PT detecting cpds. which inhibit the ability of HBV pX protein to
PT promote DNA binding to bZIP-contg. transcription factors.
XX
PS Example 1; Page 14; 47pp; English.
XX
CC The invention relates to novel methods of screening for inhibitors of
CC hepatitis B virus (HBV) replication by identifying inhibitors of the HBV
CC protein pX, which activates viral transcription by binding multiple
CC transcription factors such as proteins contg. a bZIP domain, from binding
CC to the transcription factors. This sequence was used to investigate the
CC effect of pX on the binding of the bZIP-contg. transcription factors
CC of the AP-1 family such as C/EBP, to its cognate binding site, in a gel
CC shift assay. The inhibitory cpds. may be antibodies, small organic mols.
CC or peptides.
XX
SQ Sequence 20 BP; 5 A; 5 C; 5 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCAGATTGCGCAATCTGCA 20
||||||||||||||||||||
Db 1 TGCAGATTGCGCAATCTGCA 20

RESULT 2
AAT32689/c
ID AAT32689 standard; DNA; 20 BP.
XX
AC AAT32689;
XX
DT 24-FEB-1997 (first entry)
XX
DE C/EBP binding site sequence.
XX
KW Inhibitor; hepatitis B virus; replication; protein pX; activation;
KW transcription factor; bZIP domain; AP-1; collagenase TRE binding site;
KW c-Jun; gel shift assay; ss.
XX
OS Synthetic.
XX
PN WO9617960-A2.
XX
PD 13-JUN-1996.
XX
PF 06-DEC-1995; 95WO-US16821.
XX
PR 07-DEC-1994; 94US-0351659.
XX
PA (SCRI-) SCRIPTGEN PHARM INC.
PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.
XX
PI Green MD, Lillie J, Perini G;
XX
DR WPI; 1996-287203/29.
XX
PT Identifying inhibitors of hepatitis B virus replication - by
PT detecting cpds. which inhibit the ability of HBV pX protein to
PT promote DNA binding to bZIP-contg. transcription factors.
XX
PS Example 1; Page 14; 47pp; English.
XX
CC The invention relates to novel methods of screening for inhibitors of
CC hepatitis B virus (HBV) replication by identifying inhibitors of the HBV
CC protein pX, which activates viral transcription by binding multiple
CC transcription factors such as proteins contg. a bZIP domain, from binding
CC to the transcription factors. This sequence was used to investigate the
CC effect of pX on the binding of the bZIP-contg. transcription factors
CC of the AP-1 family such as C/EBP, to its cognate binding site, in a gel
CC shift assay. The inhibitory cpds. may be antibodies, small organic mols.
CC or peptides.
XX
SQ Sequence 20 BP; 5 A; 5 C; 5 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCAGATTGCGCAATCTGCA 20
||||||||||||||||||||
Db 20 TGCAGATTGCGCAATCTGCA 1

RESULT 3
AAV46005
ID AAV46005 standard; DNA; 20 BP.
XX
AC AAV46005;
XX
DT 16-OCT-1998 (first entry)
XX
DE Immune adjuvant C/EBP.
XX
KW Immune system; adjuvant; vaccine; cancer; prophylactic; pathogenicity;
KW modulator; tolerance; regulator; helper cell; antigen; immunoglobulin;
KW Ig class; autoimmune response; T-cell; B-cell; tumour; ss.
XX
OS Class Bacteria.
XX
PN EP855184-A1.
XX
PD 29-JUL-1998.
XX
PF 23-JAN-1997; 97EP-0101019.
XX
PR 23-JAN-1997; 97EP-0101019.
XX
PA (HEEG/) HEEG K.
PA (LIPF/) LIPFORD G B.
PA (WAGN/) WAGNER H.
XX
PI Heeg K, Lipford GB, Wagner H;
XX
DR WPI; 1998-389630/34.
XX
PT Antigenic composition comprises polynucleotide fragment and antigen
PT - used as vaccine to treat or prevent e.g. cancer or pathogen
PT infections and to modulate immune response e.g. tolerance break and
PT regulation of TH1/TH2 cells
XX
PS Example 5; Page 9; 28pp; English.
XX
CC AAV45993-V46019 are fragments of bacterial polynucleotides which are
CC used as immune adjuvants for inclusion into vaccines to treat cancer and
CC for prophylaxis and/or treatment of conditions caused by pathogenic
CC micro-organisms. The polynucleotide is used for modulation of an immune
CC response and the modulation is selected from the group break of
CC tolerance, regulation of TH1/TH2 helper cell responses, switch of Ig
CC classes, treatment of autoimmune responses and induction of tolerances.
CC DNA oligomers are used to enhance the reactivity of immune cells to
CC viral, bacterial and parasitic antigens, to break tolerance in anergic T
CC and B cells e.g. against tumour antigens, as adjuvants in vaccination
CC against tumour-defined antigens and immunostimulatory substances in an
CC immune response against tumours and to suppress immune reactions of the
CC innate and acquired immune system. The composition is inexpensive and
CC stable and does not cause lethal shock, which happens with prior art
CC bacterial sequences.
XX
SQ Sequence 20 BP; 5 A; 5 C; 5 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCAGATTGCGCAATCTGCA 20

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 09:43:36 ; Search time 226.087 Seconds
(without alignments)
199.215 Million cell updates/sec

Title: US-09-355-254P-13
Perfect score: 20
Sequence: 1 tgcagattgcgcaatctgca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	17	AAT32689	C/EBP binding site
c	2	20	100.0	20	17	AAT32689	C/EBP binding site
	3	20	100.0	20	19	AAV46005	Immune adjuvant C/
c	4	20	100.0	20	19	AAV46005	Immune adjuvant C/
	5	20	100.0	20	20	AAZ25682	Transcription fact
c	6	20	100.0	20	20	AAZ25682	Transcription fact
	7	20	100.0	20	20	AAX89506	Probe for Nuclear
c	8	20	100.0	20	20	AAX89506	Probe for Nuclear
	9	20	100.0	20	21	AAA53297	Fluorescein C/EBPb
c	10	20	100.0	20	21	AAA53297	Fluorescein C/EBPb
	11	20	100.0	20	21	AAA27522	Electrophoretic mo
c	12	20	100.0	20	21	AAA27522	Electrophoretic mo
	13	20	100.0	20	21	AAA27523	Electrophoretic mo
c	14	20	100.0	20	21	AAA27523	Electrophoretic mo
	15	20	100.0	20	21	AAZ89652	Rabbit C/EBP bindi
c	16	20	100.0	20	21	AAZ89652	Rabbit C/EBP bindi
	17	20	100.0	20	22	AAH26604	C/EPB oligonucleot
c	18	20	100.0	20	22	AAH26604	C/EPB oligonucleot
	19	20	100.0	20	22	AAF87561	Consensus binding
c	20	20	100.0	20	22	AAF87561	Consensus binding
	21	20	100.0	20	24	AAL39242	Murine Toll-like r
c	22	20	100.0	20	24	AAL39242	Murine Toll-like r
	23	15.8	79.0	110	21	AAC13192	Human secreted pro
c	24	15.8	79.0	110	21	AAC13192	Human secreted pro
	25	15.4	77.0	24	24	ABA02933	Recognition site o
c	26	15.4	77.0	24	24	ABA02933	Recognition site o
	27	15.2	76.0	235	22	ABA13061	Human nervous syst
c	28	15.2	76.0	235	22	ABA13061	Human nervous syst
	29	15.2	76.0	426	24	ABK73079	Bacillus lichenifo
c	30	15.2	76.0	426	24	ABK73079	Bacillus lichenifo
	31	15.2	76.0	442	22	AAH87932	Peppermint plant o
c	32	15.2	76.0	442	22	AAH87932	Peppermint plant o
	33	15.2	76.0	528	22	AAS24874	Human ovarian PCR-
c	34	15.2	76.0	528	22	AAS24874	Human ovarian PCR-
	35	15.2	76.0	528	22	AAH83513	Human ovarian tumo
c	36	15.2	76.0	528	22	AAH83513	Human ovarian tumo
	37	15.2	76.0	541	24	ABL63609	Breast cancer rela
c	38	15.2	76.0	541	24	ABL63609	Breast cancer rela
	39	15.2	76.0	541	24	ABL64023	Breast cancer rela
c	40	15.2	76.0	541	24	ABL64023	Breast cancer rela
	41	15.2	76.0	584	22	AAS24508	Human ovarian PCR-
c	42	15.2	76.0	584	22	AAS24508	Human ovarian PCR-
	43	15.2	76.0	584	22	AAH83116	Human ovarian tumo
c	44	15.2	76.0	584	22	AAH83116	Human ovarian tumo
	45	15.2	76.0	602	23	ABL10777	Drosophila melanog

ALIGNMENTS

RESULT 1
AAT32689
ID AAT32689 standard; DNA; 20 BP.
XX
AC AAT32689;
XX
DT 24-FEB-1997 (first entry)
XX
DE C/EBP binding site sequence.
XX
KW Inhibitor; hepatitis B virus; replication; protein pX; activation;
KW transcription factor; bZIP domain; AP-1; collagenase TRE binding site;
KW c-Jun; gel shift assay; ss.
XX
OS Synthetic.
XX
PN WO9617960-A2.
XX
PD 13-JUN-1996.
XX
PF 06-DEC-1995; 95WO-US16821.
XX
PR 07-DEC-1994; 94US-0351659.
XX
PA (SCRI-) SCRIPTGEN PHARM INC.
PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.
XX
PI Green MD, Lillie J, Perini G;
XX
DR WPI; 1996-287203/29.
XX
PT Identifying inhibitors of hepatitis B virus replication - by

CC sequences of 1032, 1050 or 1032 amino acids as given in specification, or
CC their fragments, where TLR9, TLR7 and TLR8 polypeptides or their
CC fragments have an amino acid sequence which is identical to human TLR9,
CC TLR7 or TLR8 polypeptides or their fragment except for at least one amino
CC acid of a murine TLR polypeptide. The isolated nucleic acids of the
CC invention are useful for inhibiting TLR9 signalling activity in a cell.
CC TLR7, TLR8 and TLR9 polypeptides are useful for identifying nucleic acid
CC molecules which interact with a TLR polypeptide or its fragment. The
CC TLR7, TLR8 or TLR9 polypeptides are also useful for identifying ISNA. The
CC TLR7, TLR8 and TLR9 polypeptides are also useful for comparing TLR9
CC signalling activity of a test compound (that is not a nucleic acid, and
CC is a polypeptide or a part of a combinatorial library of compounds) with
CC an ISNA. The TLR7, TLR8 and TLR9 polypeptides are also useful for
CC identifying species specificity of an ISNA. The isolated nucleic acids of
CC the invention are useful as probes or primers. This polynucleotide
CC sequence represents DNA relating to the isolated Toll-like receptors of
CC the invention.
XX
SQ Sequence 20 BP; 2 A; 5 C; 11 G; 2 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 20;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGATCGGGGCGGGGCGAGC 20
||||||||||||||||||||
Db 1 TCGATCGGGGCGGGGCGAGC 20

RESULT 4
AAT18820
ID AAT18820 standard; DNA; 22 BP.
XX
AC AAT18820;
XX
DT 17-AUG-1996 (first entry)
XX
DE SP1 motif.
XX
KW NF-X1; transcription factor; major histocompatibility complex; MHC;
KW allergy; HLA-DRA; ds.
XX
OS Homo sapiens.
XX
PN WO9612823-A1.
XX
PD 02-MAY-1996.
XX
PF 20-OCT-1995; 95WO-US12749.
XX
PR 21-OCT-1994; 94US-0327832.
XX
PA (HARD) HARVARD COLLEGE.
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Ono SJ, Strominger JL;
XX
DR WPI; 1996-230621/23.
XX
PT Transcription factor, NF-X1 and DNA encoding it - used in regulation
PT of MHC class II expression and in treatment of allergic disease
XX
PS Example 4; Page 42; 93pp; English.
XX
CC Recombinant transcription factor NF-X1 (see AAR94957) forms a
CC specific complex with the HLA-DRA X1 box oligonucleotide (AAT18817)
CC which is competed for by 100-fold excess cold, double-stranded
CC oligonucleotides containing the analagous regions from the HLA-DRB,
CC -DPA, -DPB, -DQA and -DQB promoters, but not by HLA-DRA Y-box
CC (AAT18818), S-box (AAT18819), SP1 (AAT18820) or the interferon-beta gene
CC positive-regulatory domain II element (AAT18821). It is concluded
CC that NF-X1 binds sequence-specifically with all human class II
CC major histocompatibility X1 boxes (see also AAT18812).
XX
SQ Sequence 22 BP; 3 A; 5 C; 11 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGATCGGGGCGGGGCGAGC 20
||||||||||||||||||||
Db 3 TCGATCGGGGCGGGGCGAGC 22

RESULT 5
AAT77128
ID AAT77128 standard; DNA; 22 BP.
XX
AC AAT77128;
XX
DT 07-DEC-1997 (first entry)
XX
DE Sp1 consensus.
XX
KW 17-Beta-hydroxysteroid dehydrogenase type I; HSD17B1; human;
KW promoter; ds.
XX
OS Synthetic.
XX
PN WO9720942-A1.
XX
PD 12-JUN-1997.
XX
PF 04-DEC-1996; 96WO-FI00647.
XX
PR 05-DEC-1995; 95US-0007976.
XX
PA (OIKA/) OIKARINEN J A.
PA (PELT/) PELTOKETO E H.
PA (PIAO/) PIAO Y.
PA (VIHK/) VIHKO R K.
XX
PI Oikarinen JA, Peltoketo EH, Piao Y, Vihko RK;
XX
DR WPI; 1997-319788/29.
XX
PT Human 17-beta-hydroxysteroid dehydrogenase type I (HSD17B1)
PT transcription regulatory elements - used for identifying agents
PT which can up- or down-regulate HSD17B1 expression to increase or
PT decrease oestrogen production
XX
PS Example 6; Page 38; 69pp; English.
XX
CC This oligonucleotide comprises a consensus sequence for Sp1
CC binding sites. It was used with oligonucleotides (see
CC AAT77122-24 and AAT77126-27) based on the promoter region of the human
CC 17-beta-hydroxysteroid dehydrogenase type I (HSD17B1) gene (see
CC AAT77112), and with an AP-2 consensus oligonucleotide (see AAT77125),
CC in the detailed characterisation of the HSD17B1 promoter, and to
CC examine the role of Sp1 and AP-2 binding sites in promoter
CC function.
XX
SQ Sequence 22 BP; 3 A; 5 C; 11 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 18; Length 22;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGATCGGGGCGGGGCGAGC 20
||||||||||||||||||||
Db 3 TCGATCGGGGCGGGGCGAGC 22

RESULT 6
AAX76050

LIPFORD GRAYSON B (DE); HEEG KLAUS (DE)
ATURES Location/Qualifiers
source 1. .20
/organism="unidentified"
/db_xref="taxon:32644"
SE COUNT 5 a 3 c 4 g 8 t
IGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TATGCATATTCCTGTAAGTG 20
||||||||||||||||||||
1 TATGCATATTCCTGTAAGTG 20

SULT 2
0883
CUS A90883 20 bp DNA linear PAT 22-JAN-2000
FINITION Sequence 18 from Patent EP0855184.
CESSION A90883
RSION A90883.1 GI:6739312
YWORDS .
URCE unidentified.
ORGANISM unidentified
unclassified.
FERENCE 1 (bases 1 to 20)
AUTHORS Heeg,K.P. and Lipford,G.B.
TITLE Pharmaceutical composition comprising a polynucleotide and an antigen especially for vaccination
JOURNAL Patent: EP 0855184-A 18 29-JUL-1998;
HEEG KLAUS PROF DR (DE); LIPFORD GRAYSON B DR (DE)
ATURES Location/Qualifiers
source 1. .20
/organism="unidentified"
/db_xref="taxon:32644"
SE COUNT 5 a 3 c 4 g 8 t
IGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TATGCATATTCCTGTAAGTG 20
||||||||||||||||||||
1 TATGCATATTCCTGTAAGTG 20

SULT 3
455591
CUS AX455591 20 bp DNA linear PAT 06-JUL-2002
FINITION Sequence 68 from Patent WO0222809.
CESSION AX455591
RSION AX455591.1 GI:21714659
YWORDS .
URCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
FERENCE 1
AUTHORS Bauer,S., Lipford,G. and Wagner,H.
TITLE Process for high throughput screening of cpg-based immuno-agonist/antagonist
JOURNAL Patent: WO 0222809-A 68 21-MAR-2002;
Coley Pharmaceutical GmbH (DE)
ATURES Location/Qualifiers
source 1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic oligonucleotide"
3 c 4 g 8 t

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TATGCATATTCCTGTAAGTG 20
||||||||||||||||||||
Db 1 TATGCATATTCCTGTAAGTG 20

RESULT 4
MUSLY6A2
LOCUS MUSLY6A2 4524 bp DNA linear ROD 27-APR-1993
DEFINITION Mus musculus Ly-6A.2 alloantigen gene, complete cds.
ACCESSION M73552
VERSION M73552.1 GI:198925
KEYWORDS alloantigen; phosphatidylinositol-anchored cell surface protein.
SOURCE Mus musculus (strain C57BL/6) (library: EMBL-3 C57Bl/6) adult liver DNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4524)
AUTHORS Stanford,W.L., Bruyns,E. and Snodgrass,H.R.
TITLE The isolation and sequence of the chromosomal gene and regulatory regions of Ly-6A.2
JOURNAL Immunogenetics 35 (6), 408-411 (1992)
MEDLINE 92250126
PUBMED 1315719
FEATURES Location/Qualifiers
source 1. .4524
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/map="chromosome XV, Ly-6"
/haplotype="H-2b"
/tissue_type="liver"
/dev_stage="adult"
/germline
/tissue_lib="EMBL-3 C57Bl/6"
CAAT_signal 1485. .1489
/gene="Ly-6Ab"
mRNA join(1611. .1649,1895. .1976,2943. .3059,4032. .4524)
/partial
/gene="Ly-6Ab"
/product="Ly-6A.2 alloantigen"
exon 1611. .1649
/gene="Ly-6Ab"
/number=1
intron 1650. .1894
/gene="Ly-6Ab"
/number=1
exon 1895. .1976
/gene="Ly-6Ab"
/number=2
gene join(1907. .1976,2943. .3059,4032. .4249)
/gene="Ly-6Ab"
CDS join(1907. .1976,2943. .3059,4032. .4249)
/gene="Ly-6Ab"
/codon_start=1
/product="Ly-6A.2 alloantigen"
/protein_id="AAA39465.1"
/db_xref="GI:198926"
/translation="MDTSHTTKSCLLILLVALLCAERAQGLECYQCYGVPFETSCPSI
TCPYPDGVCVTQEAAAVIVDSQTRKVKNNLCLPICPPNIESMEILGTKVNVKTSCCQED
LCNVAVPNGGSTWTMAGVLLFSLSSVLLQTLL"
intron 1977. .2942
/gene="Ly-6Ab"
/number=2
exon 2943. .3059
/gene="Ly-6Ab"
/number=3
intron 3060. .4031
/gene="Ly-6Ab"

```
                     /number=3
     exon            4032. .4524
                     /partial
                     /gene="Ly-6Ab"
                     /number=4
BASE COUNT     1103 a   1031 c   1149 g   1241 t
ORIGIN      Chromosome XV.

  Query Match           100.0%;  Score 20;  DB 10;  Length 4524;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TATGCATATTCCTGTAAGTG 20
              ||||||||||||||||||||
Db        375 TATGCATATTCCTGTAAGTG 394


RESULT 5
MUSLY6A
LOCUS       MUSLY6A                 4524 bp    DNA     linear   ROD 28-JUL-1999
DEFINITION  Mus musculus Ly-6A.2 (Ly-6Ab) gene, complete cds.
ACCESSION   M74013
VERSION     M74013.1  GI:198923
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4524)
  AUTHORS   Stanford,W.L., Bruyns,E. and Snodgrass,H.R.
  JOURNAL   Unpublished
FEATURES             Location/Qualifiers
     source          1. .4524
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="15"
     gene            1611. .4524
                     /gene="Ly-6Ab"
     mRNA            join(1611. .1649,1895. .1976,2943. .3059,4032. .4524)
                     /gene="Ly-6Ab"
     CDS             join(1907. .1976,2943. .3059,4032. .4249)
                     /gene="Ly-6Ab"
                     /codon_start=1
                     /product="Ly-6A.2"
                     /protein_id="AAA39464.1"
                     /db_xref="GI:198924"
                     /translation="MDTSHTTKSCLLILLVALLCAERAQGLECYQCYGVPFETSCPSI
                     TCPYPDGVCVTQEAAVIVDSQTRKVKNNLCLPICPPNIESMEILGTKVNVKTSCCQED
                     LCNVAVPNGGSTWTMAGVLLFSLSSVLLQTLL"
BASE COUNT     1103 a   1031 c   1149 g   1241 t
ORIGIN

  Query Match           100.0%;  Score 20;  DB 10;  Length 4524;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TATGCATATTCCTGTAAGTG 20
              ||||||||||||||||||||
Db        375 TATGCATATTCCTGTAAGTG 394


RESULT 6
MUSLY6EA
LOCUS       MUSLY6EA                6249 bp    DNA     linear   ROD 27-APR-1993
DEFINITION  Mouse Ly-6E/A gene, complete cds.
ACCESSION   M37707
VERSION     M37707.1  GI:198929
KEYWORDS    differentiation antigen; interferon inducible antigen.
SOURCE      Mouse (strain BALB/c) DNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 6249)
  AUTHORS   Khan,K.D., Lindwall,G., Maher,S.E. and Bothwell,A.L.
  TITLE     Characterization of promoter elements of an interferon-inducible
            Ly-6E/A differentiation antigen, which is expressed on activated T
            cells and hematopoietic stem cells
  JOURNAL   Mol. Cell. Biol. 10 (10), 5150-5159 (1990)
  MEDLINE   90377204
   PUBMED   1697928
COMMENT     Draft entry and computer-readable sequence for [Mol. Cell. Biol.
            (1990) In press] kindly submitted
            by A.L.M.Bothwell, 13-AUG-1990.
FEATURES             Location/Qualifiers
     source          1. .6249
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
     gene            join(3206. .3258,3474. .3571,4533. .4649,5649. .>6249)
                     /gene="Ly-6E/A"
     mRNA            join(3206. .3258,3474. .3571,4533. .4649,5649. .>6249)
                     /gene="Ly-6E/A"
                     /product="differentiation antigen Ly-6E/A"
     exon            3206. .3258
                     /gene="Ly-6E/A"
                     /number=1
     intron          3259. .3473
                     /gene="Ly-6E/A"
                     /number=1
     exon            3474. .3571
                     /gene="Ly-6E/A"
                     /number=2
     CDS             join(3502. .3571,4533. .4649,5649. .5866)
                     /gene="Ly-6E/A"
                     /codon_start=1
                     /product="differentiation antigen Ly-6E/A"
                     /protein_id="AAA39467.1"
                     /db_xref="GI:198930"
                     /translation="MDTSHTTKSCLLILLVALLCAERAQGLECYQCYGVPFETSCPSI
                     TCPYPDGVCVTQEAAVIVDSQTRKVKNNLCLPICPPNIESMEILGTKVNVKTSCCQED
                     LCNVAVPNGGSTWTMAGVLLFSLSSVLLQTLL"
     mat_peptide     join(4541. .4649,5649. .5863)
                     /gene="Ly-6E/A"
                     /product="differentiation antigen Ly-6E/A"
     intron          3572. .4532
                     /gene="Ly-6E/A"
                     /number=2
     exon            4533. .4649
                     /gene="Ly-6E/A"
                     /number=3
     intron          4650. .5648
                     /gene="Ly-6E/A"
                     /number=3
     exon            5649. .>6249
                     /gene="Ly-6E/A"
                     /number=4
BASE COUNT     1592 a   1419 c   1572 g   1666 t
ORIGIN

  Query Match           100.0%;  Score 20;  DB 10;  Length 6249;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TATGCATATTCCTGTAAGTG 20
              ||||||||||||||||||||
Db       1973 TATGCATATTCCTGTAAGTG 1992


RESULT 7
AC117802/c
LOCUS       AC117802              161084 bp    DNA     linear   HTG 09-AUG-2002
DEFINITION  Mus musculus clone RP24-560I8, WORKING DRAFT SEQUENCE, 6 unordered
            pieces.
ACCESSION   AC117802
VERSION     AC117802.2  GI:22165187
```

THIS PAGE BLANK (USP

```
LOCUS       HUMIRF2A                585 bp    DNA     linear   PRI 29-MAY-2002
DEFINITION  Human gene for interferon regulatory factor-2 (IRF-2), exon 1.
ACCESSION   D14082
VERSION     D14082.1  GI:468933
KEYWORDS    IRF-2; interferon regulatory factor-2; transcription factor.
SOURCE      Homo sapiens DNA, clone_lib:Dr.Tom Maniatis's library.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Itoh,S., Harada,H., Fujita,T., Mimura,T. and Taniguchi,T.
  TITLE     Sequence of a cDNA coding for human IRF-2
  JOURNAL   Nucleic Acids Res. 17 (20), 8372 (1989)
  MEDLINE   90045964
REFERENCE   2  (bases 1 to 585)
  AUTHORS   Harada,H., Takahashi,E., Itoh,S., Harada,K., Hori,T.A. and
            Taniguchi,T.
  TITLE     Structure and regulation of the human interferon regulatory factor
            1 (IRF-1) and IRF-2 genes: implications for a gene network in the
            interferon system
  JOURNAL   Mol. Cell. Biol. 14 (2), 1500-1509 (1994)
  MEDLINE   94119101
REFERENCE   3  (bases 1 to 585)
  AUTHORS   Harada,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JAN-1993) Hisashi Harada, Osaka University, Institute
            for Molecular and Cellular Bio; 1-3,Yamada-oka, Suita, Osaka Pref.
            565, Japan (Tel:06-877-5289, Fax:06-878-9846)
COMMENT     Submitted (20-JAN-1993) to DDBJ by:
            Hisashi Harada
            Institute for Molecular and
            Cellular Biology, Osaka University
            1-3 Yamadaoka
            Suita-shi, Osaka 565
            Japan
            Phone:  06-877-5289
            Fax:    06-878-9846.
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="4"
                     /map="4q35.1."
                     /clone_lib="Dr.Tom Maniatis's library"
     exon            415. .585
                     /product="IRF-2"
                     /number=1
     misc_feature    415
                     /note="major cap site"
                     /evidence=experimental
BASE COUNT      102 a    184 c    185 g    114 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 9;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAGCGAAAATGAAATTGACT 20
              ||||||||||||||||||||
Db        153 AAGCGAAAATGAAATTGACT 172


RESULT 7
HUMIRF2A
LOCUS       HUMIRF2A                1565 bp    DNA     linear   PRI 06-JAN-1995
DEFINITION  Homo sapiens interferon regulatory factor 2 (IRF2) gene, 5' flank.
ACCESSION   L24442
VERSION     L24442.1  GI:438637
KEYWORDS    interferon regulatory factor 2.
SOURCE      Homo sapiens placenta DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1565)
  AUTHORS   Cha,Y. and Deisseroth,A.B.
  TITLE     Human interferon regulatory factor 2 gene. Intron-exon organization
            and functional analysis of 5'-flanking region
  JOURNAL   J. Biol. Chem. 269 (7), 5279-5287 (1994)
  MEDLINE   94148994
  PUBMED    8106512
FEATURES             Location/Qualifiers
     source          1. .1565
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="Unassigned"
                     /cell_type="leukocyte"
                     /tissue_type="placenta"
     protein_bind    complement(1119. .1124)
                     /note="putative"
                     /bound_moiety="Sp1"
     gene            1278. .1565
                     /gene="IRF2"
     CAAT_signal     1278. .1282
                     /gene="IRF2"
                     /note="G00-127-270; putative"
     protein_bind    1386. .1391
                     /gene="IRF2"
                     /note="putative"
                     /bound_moiety="Sp1"
     protein_bind    1400. .1405
                     /gene="IRF2"
                     /note="putative"
                     /bound_moiety="Sp1"
     protein_bind    1441. .1446
                     /gene="IRF2"
                     /note="putative"
                     /bound_moiety="Sp1"
     protein_bind    1455. .1464
                     /gene="IRF2"
                     /note="putative"
                     /bound_moiety="NF kappa B"
     mRNA            1555. .>1565
                     /gene="IRF2"
                     /note="5' untranslated region; G00-127-270"
                     /evidence=experimental
BASE COUNT      340 a    429 c    424 g    372 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 9;  Length 1565;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAGCGAAAATGAAATTGACT 20
              ||||||||||||||||||||
Db       1226 AAGCGAAAATGAAATTGACT 1245


RESULT 8
AC099343/c
LOCUS       AC099343               196216 bp    DNA     linear   PRI 22-FEB-2002
DEFINITION  Homo sapiens BAC clone RP11-326I11 from 4, complete sequence.
ACCESSION   AC099343 AC009590
VERSION     AC099343.3  GI:18543145
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 196216)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
  PUBMED    9847074
REFERENCE   2  (bases 1 to 196216)
  AUTHORS   Levy,A., Haakenson,W. and Spalding,L.
```

THIS PAGE BLANK (USPTO)

```
          16/10, Moscow V-437, 117871 GSP7, Russia
FEATURES             Location/Qualifiers
     source          1. .4740
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="leukocyte"
                     /tissue_type="blood"
                     /clone_lib="human genome library"
     gene            2124. .4176
                     /gene="IL13"
     CDS             join(2124. .2255,3312. .3365,3618. .3722,4072. .4176)
                     /gene="IL13"
                     /codon_start=1
                     /product="interleukin 13"
                     /protein_id="AAA83738.1"
                     /db_xref="GI:1127548"
                     /translation="MALLLTTVIALTCLGGFASPGPVPPSTALRELIEELVNITQNQK
                     RPLCNGSMVWSINLTAGMYCAALESLINVSGCSAIEKTQRMLSGFCPHKVSAGFSSLH
                     VRDTKIEVAQFVKDLLLHLKKLFREGRFN"
BASE COUNT     1023 a    1444 c    1305 g     968 t
ORIGIN

  Query Match           100.0%;  Score 18;  DB 9;  Length 4740;
  Best Local Similarity  100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGAATGACGTTCCCTGTG 18
              ||||||||||||||||||
Db        824 GGAATGACGTTCCCTGTG 841


RESULT 9
HSU31120
LOCUS       HSU31120                5670 bp    DNA     linear   PRI 05-JUN-1996
DEFINITION  Human interleukin-13 (IL-13) precursor gene, complete cds.
ACCESSION   U31120
VERSION     U31120.1  GI:1045451
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5670)
  AUTHORS   Dolganov,G., Bort,S., Lovett,M., Burr,J., Schubert,L., Short,D.,
            McGurn,M., Gibson,C. and Lewis,D.B.
  TITLE     Coexpression of the interleukin-13 and interleukin-4 genes
            correlates with their physical linkage in the cytokine gene cluster
            on human chromosome 5q23-31
  JOURNAL   Blood 87 (8), 3316-3326 (1996)
  MEDLINE   96184791
   PUBMED   8605348
REFERENCE   2  (bases 1 to 5670)
  AUTHORS   Dolganov,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-1995) Gregory M. Dolganov, Human Genetics,
            Genelabs, Inc., 505 Penobscot, Redwood City, CA 94063, USA
FEATURES             Location/Qualifiers
     source          1. .5670
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q23-31"
                     /clone="58D9"
     repeat_region   1360. .1490
                     /rpt_family="Alu-repeat, left monomer"
                     /rpt_type=other
     gene            2158. .5095
                     /gene="IL-13"
                     join(2158. .2345,3403. .3456,3709. .3813,4160. .5095)
                     /gene="IL-13"
                     /product="interleukin-13 precursor"
                     2158. .2345
                     /gene="IL-13"
     CDS             join(2214. .2345,3403. .3456,3709. .3813,4160. .4267)
                     /gene="IL-13"
                     /codon_start=1
                     /product="interleukin-13 precursor"
                     /protein_id="AAB01681.1"
                     /db_xref="GI:1045452"
                     /translation="MALLLTTVIALTCLGGFASPGPVPPSTALRELIEELVNITQNQK
                     APLCNGSMVWSINLTAGMYCAALESLINVSGCSAIEKTQRMLSGFCPHKVSAGQFSSL
                     HVRDTKIEVAQFVKDLLLHLKKLFREGRFN"
     sig_peptide     2214. .2273
                     /gene="IL-13"
     mat_peptide     join(2274. .2345,3403. .3456,3709. .3813,4160. .4264)
                     /gene="IL-13"
                     /product="interleukin-13"
                     /function="cytokine"
                     /evidence=experimental
     intron          2346. .3402
                     /gene="IL-13"
     exon            3403. .3456
                     /gene="IL-13"
     intron          3457. .3708
                     /gene="IL-13"
     exon            3709. .3813
                     /gene="IL-13"
     intron          3814. .4159
                     /gene="IL-13"
     exon            4160. .5095
                     /gene="IL-13"
     repeat_region   4831. .4878
                     /rpt_type=direct
     polyA_site      5076. .5081
                     /gene="IL-13"
     repeat_region   5522. .5567
BASE COUNT     1243 a    1676 c    1582 g    1169 t
ORIGIN

  Query Match           100.0%;  Score 18;  DB 9;  Length 5670;
  Best Local Similarity  100.0%;  Pred. No. 16;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGAATGACGTTCCCTGTG 18
              ||||||||||||||||||
Db        826 GGAATGACGTTCCCTGTG 843


RESULT 10
AF377331
LOCUS       AF377331                6919 bp    DNA     linear   PRI 01-JUN-2001
DEFINITION  Homo sapiens interleukin 13 (IL13) gene, complete cds.
ACCESSION   AF377331
VERSION     AF377331.2  GI:14278714
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6919)
  AUTHORS   Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAY-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
REFERENCE   2  (bases 1 to 6919)
  AUTHORS   Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
            Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-2001) Molecular Biotechnology, University of
            Washington, 1705 NE Pacific, Seattle, WA 98195, USA
   REMARK   Sequence update by submitter
COMMENT     On Jun 1, 2001 this sequence version replaced gi:14091715.
            To cite this work please use:  SeattleSNPs. NHLBI Program for
            Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.mbt.washington.edu).
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 10:58:25 ; Search time 30.2174 Seconds
(without alignments)
202.980 Million cell updates/sec

Title: US-09-355-254F-12
Perfect score: 20
Sequence: 1 tcgatcggggcggggcgagc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	22	2	US-08-327-832-12	Sequence 12, Appl
	2	20	100.0	22	2	US-08-828-584-12	Sequence 12, Appl
	3	20	100.0	22	5	PCT-US94-05659-16	Sequence 16, Appl
	4	19	95.0	46	1	US-08-122-433-12	Sequence 12, Appl
	5	19	95.0	46	1	US-08-122-433-13	Sequence 13, Appl
c	6	18	90.0	46	1	US-08-122-433-12	Sequence 12, Appl
c	7	18	90.0	46	1	US-08-122-433-13	Sequence 13, Appl
c	8	16.8	84.0	2875	3	US-08-458-434A-4	Sequence 4, Appli
c	9	16.8	84.0	15144	3	US-08-458-434A-6	Sequence 6, Appli
c	10	16.4	82.0	379	1	US-08-145-617-5	Sequence 5, Appli
c	11	15.8	79.0	31	1	US-08-153-563-4	Sequence 4, Appli
c	12	15.8	79.0	31	2	US-09-038-227-9	Sequence 9, Appli
c	13	15.8	79.0	31	2	US-08-460-507-4	Sequence 4, Appli
	14	15.8	79.0	944	2	US-08-786-606-4	Sequence 4, Appli
	15	15.8	79.0	1112	2	US-08-933-750C-97	Sequence 97, Appl
	16	15.8	79.0	1112	3	US-09-234-613-97	Sequence 97, Appl
c	17	15.8	79.0	1919	1	US-07-991-587A-1	Sequence 1, Appli
c	18	15.8	79.0	1919	1	US-08-309-985-1	Sequence 1, Appli
c	19	15.8	79.0	3314	1	US-07-973-324A-5	Sequence 5, Appli
c	20	15.8	79.0	3314	1	US-08-343-380-5	Sequence 5, Appli
c	21	15.8	79.0	3314	4	US-09-072-435-5	Sequence 5, Appli
c	22	15.8	79.0	3314	4	US-09-072-917A-5	Sequence 5, Appli
c	23	15.2	76.0	80	3	US-09-039-555B-4	Sequence 4, Appli
c	24	15.2	76.0	730	2	US-08-743-637B-11	Sequence 11, Appl
c	25	15.2	76.0	730	3	US-08-526-840B-11	Sequence 11, Appl
c	26	15.2	76.0	1007	4	US-08-836-500A-1	Sequence 1, Appli
c	27	15.2	76.0	1008	3	US-08-721-979A-13	Sequence 13, Appl
c	28	15.2	76.0	1008	4	US-09-654-289-13	Sequence 13, Appl
	29	15.2	76.0	9704	4	US-09-814-951A-3	Sequence 3, Appli
	30	15.2	76.0	44377	2	US-08-804-227C-7	Sequence 7, Appli
	31	15.2	76.0	44377	2	US-08-804-198-1	Sequence 1, Appli
c	32	15.2	76.0	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	33	15.2	76.0	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	34	14.8	74.0	645	4	US-08-998-416-114	Sequence 114, App
	35	14.8	74.0	4060	1	US-08-164-292B-1	Sequence 1, Appli
	36	14.8	74.0	4060	1	US-08-164-292B-3	Sequence 3, Appli
	37	14.8	74.0	4060	1	US-08-164-292B-5	Sequence 5, Appli
	38	14.8	74.0	4060	1	US-08-164-292B-7	Sequence 7, Appli
	39	14.8	74.0	4060	3	US-08-845-623-1	Sequence 1, Appli
	40	14.8	74.0	4060	3	US-08-845-623-3	Sequence 3, Appli
	41	14.8	74.0	4060	3	US-08-845-623-5	Sequence 5, Appli
	42	14.8	74.0	4060	3	US-08-845-623-7	Sequence 7, Appli
	43	14.8	74.0	4060	3	US-08-815-927-1	Sequence 1, Appli
	44	14.8	74.0	4060	3	US-08-815-927-3	Sequence 3, Appli
	45	14.8	74.0	4060	3	US-08-815-927-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-327-832-12
; Sequence 12, Application US/08327832
; Patent No. 5840832
; GENERAL INFORMATION:
; APPLICANT: Ono, Santa J.
; APPLICANT: Strominger, Jack L.
; TITLE OF INVENTION: Transcription Factor Regulating MHC
; TITLE OF INVENTION: Expression, cDNA and Genomic Clones Encoding Same and
; TITLE OF INVENTION: Retroviral Expression Contructs Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/327,832
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.46362
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 20-2 508-9153
; TELEFAX: 202 508-9299
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
US-08-327-832-12

Query Match 100.0%; Score 20; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGATCGGGGCGGGGCGAGC 20

```
                    ||||||||||||||||||||
Db          3 TCGATCGGGGCGGGGCGAGC 22


RESULT 2
US-08-828-584-12
; Sequence 12, Application US/08828584
; Patent No. 5908762
;  GENERAL INFORMATION:
;    APPLICANT:  Ono, Santa J.
;    APPLICANT:  Strominger, Jack L.
;    TITLE OF INVENTION:  Transcription Factor Regulating MHC
;    TITLE OF INVENTION:  Expression, cDNA and Genomic Clones Encoding Same and
;    TITLE OF INVENTION:  Retroviral Expression Contructs Thereof
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner, Birch, McKie & Beckett
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington, D.C.
;      STATE:  District of Columbia
;      COUNTRY:  U.S.A.
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/828,584
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Posorske, Laurence H.
;      REGISTRATION NUMBER:  34,698
;      REFERENCE/DOCKET NUMBER:  1107.46362
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  20-2 508-9153
;      TELEFAX:  202 508-9299
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  homo sapiens
US-08-828-584-12

  Query Match             100.0%;  Score 20;  DB 2;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCGATCGGGGCGGGGCGAGC 20
                ||||||||||||||||||||
Db          3 TCGATCGGGGCGGGGCGAGC 22


RESULT 3
PCT-US94-05659-16
; Sequence 16, Application PC/TUS9405659
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  TNF' RESPONSIVE ELEMENT, TNF'-INDUCED DNA-BINDING
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/05659
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  FDC93-01 FF
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US94-05659-16

  Query Match             100.0%;  Score 20;  DB 5;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCGATCGGGGCGGGGCGAGC 20
                ||||||||||||||||||||
Db          3 TCGATCGGGGCGGGGCGAGC 22


RESULT 4
US-08-122-433-12
; Sequence 12, Application US/08122433
; Patent No. 5683985
;  GENERAL INFORMATION:
;    APPLICANT:  Chu, Barbara C.F.
;    APPLICANT:  Orgel, Leslie
;    TITLE OF INVENTION:  OLIGODEOXYNUCLEOTIDES AND
;    TITLE OF INVENTION:  OLIGONUCLEOTIDES USEFUL AS DECOYS FOR PROTEINS WHICH
;    TITLE OF INVENTION:  SELECTIVELY BIND TO DEFINED DNA SEQUENCES
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/122,433
;      FILING DATE:  22-SEP-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/687,337
;      FILING DATE:  18-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P31 9308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-1995
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
US-08-122-433-12

  Query Match             95.0%;  Score 19;  DB 1;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 7.8;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCGATCGGGGCGGGGCGAG 19
              ||||||||||||||||||||
Db         28 TCGATCGGGGCGGGGCGAG 46


RESULT 5
US-08-122-433-13
; Sequence 13, Application US/08122433
; Patent No. 5683985
;  GENERAL INFORMATION:
;    APPLICANT:  Chu, Barbara C.F.
;    APPLICANT:  Orgel, Leslie
;    TITLE OF INVENTION:  OLIGODEOXYNUCLEOTIDES AND
;    TITLE OF INVENTION:  OLIGONUCLEOTIDES USEFUL AS DECOYS FOR PROTEINS WHICH
;    TITLE OF INVENTION:  SELECTIVELY BIND TO DEFINED DNA SEQUENCES
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/122,433
;      FILING DATE:  22-SEP-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/687,337
;      FILING DATE:  18-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P31 9308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-1995
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  other nucleic acid
US-08-122-433-13

  Query Match             95.0%;  Score 19;  DB 1;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 7.8;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCGATCGGGGCGGGGCGAG 19
              ||||||||||||||||||||
Db         28 TCGATCGGGGCGGGGCGAG 46


RESULT 6
US-08-122-433-12/c
; Sequence 12, Application US/08122433
; Patent No. 5683985
;  GENERAL INFORMATION:
;    APPLICANT:  Chu, Barbara C.F.
;    APPLICANT:  Orgel, Leslie
;    TITLE OF INVENTION:  OLIGODEOXYNUCLEOTIDES AND
;    TITLE OF INVENTION:  OLIGONUCLEOTIDES USEFUL AS DECOYS FOR PROTEINS WHICH
;    TITLE OF INVENTION:  SELECTIVELY BIND TO DEFINED DNA SEQUENCES
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/122,433
;      FILING DATE:  22-SEP-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/687,337
;      FILING DATE:  18-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P31 9308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-1995
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
US-08-122-433-12

  Query Match             90.0%;  Score 18;  DB 1;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GATCGGGGCGGGGCGAGC 20
              ||||||||||||||||||
Db         24 GATCGGGGCGGGGCGAGC 7


RESULT 7
US-08-122-433-13/c
; Sequence 13, Application US/08122433
; Patent No. 5683985
;  GENERAL INFORMATION:
;    APPLICANT:  Chu, Barbara C.F.
;    APPLICANT:  Orgel, Leslie
;    TITLE OF INVENTION:  OLIGODEOXYNUCLEOTIDES AND
;    TITLE OF INVENTION:  OLIGONUCLEOTIDES USEFUL AS DECOYS FOR PROTEINS WHICH
;    TITLE OF INVENTION:  SELECTIVELY BIND TO DEFINED DNA SEQUENCES
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/122,433
;      FILING DATE:  22-SEP-1993
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/687,337
;      FILING DATE:  18-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P31 9308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-1995
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  other nucleic acid
US-08-122-433-13

  Query Match             90.0%;  Score 18;  DB 1;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GATCGGGGCGGGGCGAGC 20
              ||||||||||||||||||
Db         24 GATCGGGGCGGGGCGAGC 7


RESULT 8
US-08-458-434A-4/c
; Sequence 4, Application US/08458434A
; Patent No. 6083690
;  GENERAL INFORMATION:
;    APPLICANT:  Harris Ph.D., Stephen E.
;    APPLICANT:  Mundy M.D., Gregory R.
;    APPLICANT:  Gosh-Choudhury Ph.D., Nandini
;    APPLICANT:  Feng Ph.D., Jian Q.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR IDENTIFYING
;    TITLE OF INVENTION:  OSTEOGENIC AGENTS
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  James C. Weseman, Esq.
;      STREET:  401 B. Street, Suite 1700
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,434A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weseman, James C.
;      REGISTRATION NUMBER:  30,507
;      REFERENCE/DOCKET NUMBER:  P00060US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 699-3604
;      TELEFAX:  619-236-1048
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2875 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-458-434A-4

  Query Match             84.0%;  Score 16.8;  DB 3;  Length 2875;
  Best Local Similarity   90.0%;  Pred. No. 41;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TCGATCGGGGCGGGGCGAGC 20
              | || ||||||||||||||||
Db       1805 TGGAGCGGGGCGGGGCGAGC 1786


RESULT 9
US-08-458-434A-6/c
; Sequence 6, Application US/08458434A
; Patent No. 6083690
;  GENERAL INFORMATION:
;    APPLICANT:  Harris Ph.D., Stephen E.
;    APPLICANT:  Mundy M.D., Gregory R.
;    APPLICANT:  Gosh-Choudhury Ph.D., Nandini
;    APPLICANT:  Feng Ph.D., Jian Q.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR IDENTIFYING
;    TITLE OF INVENTION:  OSTEOGENIC AGENTS
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  James C. Weseman, Esq.
;      STREET:  401 B. Street, Suite 1700
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,434A
;      FILING DATE:
;      CLASSIFICATION:  435 .
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weseman, James C.
;      REGISTRATION NUMBER:  30,507
;      REFERENCE/DOCKET NUMBER:  P00060US0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 699-3604
;      TELEFAX:  619-236-1048
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15144 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-458-434A-6

  Query Match             84.0%;  Score 16.8;  DB 3;  Length 15144;
  Best Local Similarity   90.0%;  Pred. No. 35;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TCGATCGGGGCGGGGCGAGC 20
              | || ||||||||||||||||
Db       1805 TGGAGCGGGGCGGGGCGAGC 1786


RESULT 10
US-08-145-617-5/c
; Sequence 5, Application US/08145617
; Patent No. 5766847
;  GENERAL INFORMATION:
;    APPLICANT:  Jackle, Herbert
```

```
;    APPLICANT:  Tautz, Diethard
;    TITLE OF INVENTION:  PROCESS FOR ANALYZING LENGTH
;    TITLE OF INVENTION:  POLYMORPHISMS IN DNA REGIONS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  301 N. Washington Street, P.O. Box 747
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  United States of America
;      ZIP:  22046
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/145,617
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/681,494
;      FILING DATE:  10-JUN-1991
;      APPLICATION NUMBER:  DE P3834636.2
;      FILING DATE:  11-OCT-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Svensson, Leonard R.
;      REGISTRATION NUMBER:  30,330
;      REFERENCE/DOCKET NUMBER:  147-122PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  379 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-145-617-5

  Query Match             82.0%;  Score 16.4;  DB 1;  Length 379;
  Best Local Similarity   94.4%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||
Db         64 GATTGGGGCGGGGCGAGC 47


RESULT 11
US-08-153-563-4/c
; Sequence 4, Application US/08153563
; Patent No. 5693506
;  GENERAL INFORMATION:
;    APPLICANT:  Rodriquez, Raymond L.
;    TITLE OF INVENTION:  PROCESS FOR PROTEIN PRODUCTION
;    TITLE OF INVENTION:  IN PLANTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  Steuart Street Tower, One Market Plaza
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94105-1493
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/153,563
;      FILING DATE:  16-NOV-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Kenneth A.
;      REGISTRATION NUMBER:  31,677
;      REFERENCE/DOCKET NUMBER:  2307E-515
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..31
;      OTHER INFORMATION:  /standard_name= "31 bp RAmy3E"
US-08-153-563-4

  Query Match             79.0%;  Score 15.8;  DB 1;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 1.6e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              |||||| ||||| ||||||
Db         23 CGATCGAGGCGGCGCGAGC 5


RESULT 12
US-09-038-227-9/c
; Sequence 9, Application US/09038227
; Patent No. 5917029
;  GENERAL INFORMATION:
;    APPLICANT:  Yu, Su-May
;    TITLE OF INVENTION:  SUGAR-RESPONSIVE ENHANCERS
;    TITLE OF INVENTION:  IN ALPHA-AMYLASE GENES
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/038,227
;      FILING DATE:  11-MAR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tsao, Y. Rocky
;      REGISTRATION NUMBER:  34,053
;      REFERENCE/DOCKET NUMBER:  05228/031001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  Genomic DNA
US-09-038-227-9

  Query Match             79.0%;  Score 15.8;  DB 2;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 1.6e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              |||||| ||||| ||||||
Db         23 CGATCGAGGCGGCGCGAGC 5


RESULT 13
US-08-460-507-4/c
; Sequence 4, Application US/08460507
; Patent No. 5994628
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  PROCESS FOR PROTEIN PRODUCTION
;    TITLE OF INVENTION:  IN PLANTS
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/460,507
;      FILING DATE:
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/153,563
;      FILING DATE:  16-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  2000-0452.41
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..31
;      OTHER INFORMATION:  /standard_name= "31 bp RAmy3E"
US-08-460-507-4

  Query Match             79.0%;  Score 15.8;  DB 2;  Length 31;
  Best Local Similarity   89.5%;  Pred. No. 1.6e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              |||||| ||||| ||||||
Db         23 CGATCGAGGCGGCGCGAGC 5


RESULT 14
US-08-786-606-4
; Sequence 4, Application US/08786606
; Patent No. 5861495
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Coleman, Roger
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN ZINC-BINDING
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/786,606
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy RJ
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0173 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  944 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-786-606-4

  Query Match             79.0%;  Score 15.8;  DB 2;  Length 944;
  Best Local Similarity   89.5%;  Pred. No. 1.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| |||||||||||| ||
Db        145 CGAGCGGGGCGGGGCGGGC 163


RESULT 15
US-08-933-750C-97
; Sequence 97, Application US/08933750C
; Patent No. 5932442
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  HUMAN REGULATORY MOLECULES
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
```

; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: TESTNOT07
; CLONE: 3217567
US-08-933-750C-97

Query Match 79.0%; Score 15.8; DB 2; Length 1112;
Best Local Similarity 89.5%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| |||||||||||| ||
Db        147 CGAGCGGGGCGGGGCGGGC 165
```

Search completed: June 26, 2003, 16:20:51
Job time : 35.2888 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 10:58:25 ; Search time 30.2174 Seconds
(without alignments)
202.980 Million cell updates/sec

Title: US-09-355-254F-11
Perfect score: 20
Sequence: 1 gcttgatgactcagccggaa 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	21	2	US-08-507-598-2	Sequence 2, Appli
	2	20	100.0	21	2	US-08-507-750-2	Sequence 2, Appli
	3	20	100.0	21	3	US-08-764-522A-2	Sequence 2, Appli
	4	20	100.0	21	3	US-08-764-528-2	Sequence 2, Appli
	5	20	100.0	21	3	US-08-872-859-2	Sequence 2, Appli
	6	18.4	92.0	21	1	US-08-283-591-15	Sequence 15, Appl
	7	18.4	92.0	21	1	US-08-210-880B-3	Sequence 3, Appli
c	8	18.4	92.0	21	2	US-08-632-275-1	Sequence 1, Appli
	9	18.4	92.0	21	2	US-08-771-411-3	Sequence 3, Appli
c	10	18.4	92.0	21	3	US-09-097-929-1	Sequence 1, Appli
	11	18.4	92.0	21	4	US-09-021-247-8	Sequence 8, Appli
	12	18.4	92.0	21	4	US-08-088-661F-8	Sequence 8, Appli
	13	18.4	92.0	21	4	US-08-088-661F-42	Sequence 42, Appl
	14	15.2	76.0	1400	3	US-09-191-099-7	Sequence 7, Appli
	15	15.2	76.0	1446	3	US-09-191-099-8	Sequence 8, Appli
	16	15.2	76.0	5894	3	US-08-665-259-24	Sequence 24, Appl
	17	15.2	76.0	5894	3	US-08-762-500-24	Sequence 24, Appl
	18	15.2	76.0	6525	3	US-08-762-500-74	Sequence 74, Appl
	19	14.8	74.0	80161	3	US-09-036-987A-1	Sequence 1, Appli
	20	14.8	74.0	80161	4	US-09-370-700-1	Sequence 1, Appli
c	21	14.4	72.0	903	4	US-09-457-046B-5	Sequence 5, Appli
	22	14.4	72.0	1314	4	US-08-800-682-1	Sequence 1, Appli
c	23	14.4	72.0	1338	4	US-09-457-046B-51	Sequence 51, Appl
	24	14.4	72.0	22306	4	US-09-453-702B-251	Sequence 251, App
c	25	14.4	72.0	46819	4	US-09-453-702B-72	Sequence 72, Appl
	26	14.2	71.0	400	1	US-08-301-718-1	Sequence 1, Appli
c	27	14.2	71.0	534	2	US-08-770-544-19	Sequence 19, Appl
	28	14.2	71.0	590	2	US-08-600-999-1	Sequence 1, Appli
c	29	14.2	71.0	1251	4	US-09-355-115-1	Sequence 1, Appli
c	30	14.2	71.0	1462	3	US-08-961-083-41	Sequence 41, Appl
	31	14.2	71.0	2384	1	US-07-688-352C-27	Sequence 27, Appl
	32	14.2	71.0	2384	2	US-08-474-379C-27	Sequence 27, Appl
	33	14.2	71.0	2384	3	US-09-146-249A-27	Sequence 27, Appl
	34	14.2	71.0	2384	3	US-08-206-188B-27	Sequence 27, Appl
	35	14.2	71.0	2384	5	PCT-US91-02714-25	Sequence 25, Appl
	36	14.2	71.0	15239	1	US-08-390-878-17	Sequence 17, Appl
c	37	14.2	71.0	19702	4	US-08-961-527-7	Sequence 7, Appli
	38	14.2	71.0	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	39	14.2	71.0	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	40	14.2	71.0	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	41	14.2	71.0	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	42	13.8	69.0	21	1	US-08-203-198-1	Sequence 1, Appli
	43	13.8	69.0	945	4	US-09-149-476-168	Sequence 168, App
	44	13.8	69.0	12720	1	US-08-403-866-11	Sequence 11, Appl
c	45	13.6	68.0	26	1	US-07-791-213D-68	Sequence 68, Appl

ALIGNMENTS

RESULT 1
US-08-507-598-2
; Sequence 2, Application US/08507598
; Patent No. 5834188
; GENERAL INFORMATION:
; APPLICANT: HARADA, SHUN-ICHI
; APPLICANT: SAMPATH, T. K.
; APPLICANT: RODAN, GIDEON A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING
; TITLE OF INVENTION: MORPHOGEN ANALOGS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; ADDRESSEE: THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/507,598
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)-248-7000
; TELEFAX: (617)-248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..21
; OTHER INFORMATION: /product= "AP1 SEQUENCE"
US-08-507-598-2

Query Match 100.0%; Score 20; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.058;

```
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||||||||||||||
Db          2 GCTTGATGACTCAGCCGGAA 21


RESULT 2
US-08-507-750-2
; Sequence 2, Application US/08507750
; Patent No. 5932716
;  GENERAL INFORMATION:
;    APPLICANT:  SAMPATH, T. K.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR IDENTIFYING
;    TITLE OF INVENTION:  MORPHOGEN ANALOGS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PATENT ADMINISTRATOR, TESTA, HURWITZ &
;      ADDRESSEE:  THIBEAULT
;      STREET:  53 STATE STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/507,750
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PITCHER, EDMUND R.
;      REGISTRATION NUMBER:  27,829
;      REFERENCE/DOCKET NUMBER:  CRP-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)-248-7000
;      TELEFAX:  (617)-248-7100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..21
;      OTHER INFORMATION:  /product= "AP1 SEQUENCE"
US-08-507-750-2

  Query Match           100.0%;  Score 20;  DB 2;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 0.058;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||||||||||||||
Db          2 GCTTGATGACTCAGCCGGAA 21


RESULT 3
US-08-764-522A-2
; Sequence 2, Application US/08764522A
; Patent No. 6090544
;  GENERAL INFORMATION:
;    APPLICANT:  HARADA, SHUN-ICHI
;    APPLICANT:  SAMPATH, T. K.
;    APPLICANT:  RODAN, GIDEON A.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR IDENTIFYING
;    TITLE OF INVENTION:  MORPHOGEN ANALOGS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
;      STREET:  45 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  01748
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,522A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  VITO, CHRISTINE C.
;      REGISTRATION NUMBER:  39,061
;      REFERENCE/DOCKET NUMBER:  CRP-126
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)-248-7000
;      TELEFAX:  (617)-248-7100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..21
;      OTHER INFORMATION:  /product= "AP1 SEQUENCE A"
US-08-764-522A-2

  Query Match           100.0%;  Score 20;  DB 3;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 0.058;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||||||||||||||
Db          2 GCTTGATGACTCAGCCGGAA 21


RESULT 4
US-08-764-528-2
; Sequence 2, Application US/08764528
; Patent No. 6103491
;  GENERAL INFORMATION:
;    APPLICANT:  SAMPATH, K. T.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR IDENTIFYING
;    TITLE OF INVENTION:  MORPHOGEN ANALOGS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
;      STREET:  45 SOUTH STREET
;      CITY:  HOPKINTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  01748
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,528
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  VITO, CHRISTINE C.
;      REGISTRATION NUMBER:  39,061
```

```
;      REFERENCE/DOCKET NUMBER:  CRP-127
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)-248-7000
;      TELEFAX:  (617)-248-7100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..21
;      OTHER INFORMATION:  /product= "AP1 SEQUENCE A"
US-08-764-528-2

  Query Match             100.0%;  Score 20;  DB 3;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.058;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||||||||||||||
Db          2 GCTTGATGACTCAGCCGGAA 21


RESULT 5
US-08-872-859-2
; Sequence 2, Application US/08872859
; Patent No. 6110460
;  GENERAL INFORMATION:
;    APPLICANT:  SAMPATH, T. K.
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS FOR IDENTIFYING
;    TITLE OF INVENTION:  MORPHOGEN ANALOGS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PATENT ADMINISTRATOR, TESTA, HURWITZ &
;      ADDRESSEE:  THIBEAULT
;      STREET:  53 STATE STREET
;      CITY:  BOSTON
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/872,859
;      FILING DATE:  11-JUN-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/507,750
;      FILING DATE:  26-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PITCHER, EDMUND R.
;      REGISTRATION NUMBER:  27,829
;      REFERENCE/DOCKET NUMBER:  CRP-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)-248-7000
;      TELEFAX:  (617)-248-7100
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..21
;      OTHER INFORMATION:  /product= "AP1 SEQUENCE"
US-08-872-859-2

  Query Match             100.0%;  Score 20;  DB 3;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.058;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||||||||||||||
Db          2 GCTTGATGACTCAGCCGGAA 21


RESULT 6
US-08-283-591-15
; Sequence 15, Application US/08283591
; Patent No. 5629152
;  GENERAL INFORMATION:
;    APPLICANT:  Ravikumar, Vasulinga
;    TITLE OF INVENTION:  NOVEL TRISUBSTITUTED  -LACTAMS AND
;    TITLE OF INVENTION:  OLIGO  -LACTAMAMIDES
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz and
;      ADDRESSEE:  No. 5629152ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/283,591
;      FILING DATE:  N/A
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:
;      REGISTRATION NUMBER:
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
US-08-283-591-15

  Query Match             92.0%;  Score 18.4;  DB 1;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 0.42;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||| ||||||||||||||
Db          2 GCTTGGTGACTCAGCCGGAA 21


RESULT 7
US-08-210-880B-3
; Sequence 3, Application US/08210880B
; Patent No. 5641486
;  GENERAL INFORMATION:
;    APPLICANT:  HINRICHS, STEVEN H.
;    APPLICANT:  ORTEN, DANA J.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR INHIBITING
;    TITLE OF INVENTION:  CREB/ATF1 FAMILY DETERMINED TRANSCRIPTION
```

```
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HENDERSON & STURM
;      STREET:  1125 S. 103RD ST., #330
;      CITY:  OMAHA
;      STATE:  NE
;      COUNTRY:  US
;      ZIP:  68124
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,880B
;      FILING DATE:  18-MAR-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JONDLE, ROBERT J.
;      REGISTRATION NUMBER:  33,915
;      REFERENCE/DOCKET NUMBER:  63066
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  402-398-9000
;      TELEFAX:  402-398-9005
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-210-880B-3

  Query Match             92.0%;  Score 18.4;  DB 1;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 0.42;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||| ||||||||||
Db          2 GCTTGATGAGTCAGCCGGAA 21


RESULT 8
US-08-632-275-1/c
; Sequence 1, Application US/08632275
; Patent No. 5840277
;  GENERAL INFORMATION:
;    APPLICANT:  Ghio, Andrew J.
;    APPLICANT:  Kennedy, Thomas P.
;    TITLE OF INVENTION:  Treatment of Chronic Pulmonary
;    TITLE OF INVENTION:  Inflamation
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bell, Seltzer, Park & Gibson
;      STREET:  1211 East Morehead Street
;      CITY:  Charlotte
;      STATE:  No. 5840277th Carolina
;      COUNTRY:  USA
;      ZIP:  28234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/632,275
;      FILING DATE:  15-APR-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/413,699
;      FILING DATE:  30-MAR-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lipscomb, Ernest B.
;      REGISTRATION NUMBER:  24,733
;      REFERENCE/DOCKET NUMBER:  8751-5-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  704-331-6000
;      TELEFAX:  704-334-2014
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FRAGMENT TYPE:  linear
US-08-632-275-1

  Query Match             92.0%;  Score 18.4;  DB 2;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 0.42;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||| ||||||||||
Db         20 GCTTGATGAGTCAGCCGGAA 1


RESULT 9
US-08-771-411-3
; Sequence 3, Application US/08771411
; Patent No. 5844096
;  GENERAL INFORMATION:
;    APPLICANT:  HINRICHS, STEVEN H.
;    APPLICANT:  ORTEN, DANA J.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR INHIBITING
;    TITLE OF INVENTION:  CREB/ATF1 FAMILY DETERMINED TRANSCRIPTION
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HENDERSON & STURM
;      STREET:  1125 S. 103RD ST., #330
;      CITY:  OMAHA
;      STATE:  NE
;      COUNTRY:  US
;      ZIP:  68124
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/771,411
;      FILING DATE:  20-DEC-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/210,880
;      FILING DATE:  18-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JONDLE, ROBERT J.
;      REGISTRATION NUMBER:  33,915
;      REFERENCE/DOCKET NUMBER:  63066
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  402-398-9000
;      TELEFAX:  402-398-9005
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-771-411-3

  Query Match             92.0%;  Score 18.4;  DB 2;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 0.42;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||| ||||||||||
Db          2 GCTTGATGAGTCAGCCGGAA 21


RESULT 10
US-09-097-929-1/c
; Sequence 1, Application US/09097929
; Patent No. 6024940
;  GENERAL INFORMATION:
;    APPLICANT:  Ghio, Andrew J.
;    APPLICANT:  Kennedy, Thomas P.
;    TITLE OF INVENTION:  Treatment of Chronic Pulmonary
;    TITLE OF INVENTION:  Inflamation
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bell, Seltzer, Park & Gibson
;      STREET:  1211 East Morehead Street
;      CITY:  Charlotte
;      STATE:  No. 6024940th Carolina
;      COUNTRY:  USA
;      ZIP:  28234
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/097,929
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/632,275
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lipscomb, Ernest B.
;      REGISTRATION NUMBER:  24,733
;      REFERENCE/DOCKET NUMBER:  8751-5-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  704-331-6000
;      TELEFAX:  704-334-2014
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FRAGMENT TYPE:  linear
US-09-097-929-1

  Query Match             92.0%;  Score 18.4;  DB 3;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 0.42;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||| ||||||||||
Db         20 GCTTGATGAGTCAGCCGGAA 1


RESULT 11
US-09-021-247-8
; Sequence 8, Application US/09021247
; Patent No. 6225444
;  GENERAL INFORMATION:
;    APPLICANT:  Shashoua, Victor E.
;    TITLE OF INVENTION:  NEUROPROTECTIVE PEPTIDES AND USES THEREOF
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/021,247
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Van Amsterdam, John R.
;      REGISTRATION NUMBER:  40,212
;      REFERENCE/DOCKET NUMBER:  N0260/7023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  oligonucleotide
;    HYPOTHETICAL:  NO
US-09-021-247-8

  Query Match             92.0%;  Score 18.4;  DB 4;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 0.42;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||||||| ||||||||||
Db          2 GCTTGATGAGTCAGCCGGAA 21


RESULT 12
US-08-088-661F-8
; Sequence 8, Application US/08088661F
; Patent No. 6228982
; GENERAL INFORMATION:
;  APPLICANT:  No. 6228982den, Benget
;  APPLICANT:  Wittung, Pernilla
;  APPLICANT:  Buchardt, Ole
;  APPLICANT:  Egholm, Michael
;  APPLICANT:  Nielsen, Peter E.
;  APPLICANT:  Berg, Rolf
;  TITLE OF INVENTION:  Double-Stranded Peptide Nucleic Acids
;  FILE REFERENCE:  ISIS1108
;  CURRENT APPLICATION NUMBER: US/08/088,661F
;  CURRENT FILING DATE:  1993-07-02
;  PRIOR APPLICATION NUMBER:  08/054,363
;  PRIOR FILING DATE:  1993-04-26
;  PRIOR APPLICATION NUMBER:  PCT/EP92/01219
;  PRIOR FILING DATE:  1992-05-19
;  NUMBER OF SEQ ID NOS:  42
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: No. 6228982el Sequence
US-08-088-661F-8

  Query Match             92.0%;  Score 18.4;  DB 4;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 0.42;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||| ||||||||||||||
Db          2 GCTTGGTGACTCAGCCGGAA 21
```

```
RESULT 13
US-08-088-661F-42
; Sequence 42, Application US/08088661F
; Patent No. 6228982
; GENERAL INFORMATION:
;  APPLICANT:  No. 6228982den, Benget
;  APPLICANT:  Wittung, Pernilla
;  APPLICANT:  Buchardt, Ole
;  APPLICANT:  Egholm, Michael
;  APPLICANT:  Nielsen, Peter E.
;  APPLICANT:  Berg, Rolf
;  TITLE OF INVENTION:  Double-Stranded Peptide Nucleic Acids
;  FILE REFERENCE:  ISIS1108
;  CURRENT APPLICATION NUMBER: US/08/088,661F
;  CURRENT FILING DATE:  1993-07-02
;  PRIOR APPLICATION NUMBER:  08/054,363
;  PRIOR FILING DATE:  1993-04-26
;  PRIOR APPLICATION NUMBER:  PCT/EP92/01219
;  PRIOR FILING DATE:  1992-05-19
;  NUMBER OF SEQ ID NOS:  42
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 42
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: No. 6228982el Sequence
US-08-088-661F-42

  Query Match             92.0%;  Score 18.4;  DB 4;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 0.42;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              ||||| ||||||||||||||
Db          2 GCTTGGTGACTCAGCCGGAA 21


RESULT 14
US-09-191-099-7
; Sequence 7, Application US/09191099
; Patent No. 6096323
; GENERAL INFORMATION:
;  APPLICANT: Walker, Richard L.
;  APPLICANT: Read, Deryck H.
;  APPLICANT: Hird, David W.
;  APPLICANT: LeFebvre, Rance B.
;  APPLICANT: Berry, Steven L.
;  APPLICANT: Cullor, James S.
;  APPLICANT: Lefler, Hank M.
;  APPLICANT: The Regents of the University of California
;  TITLE OF INVENTION: Vaccine Against Papillomatous Digital Dermatitis (PDD)
;  FILE REFERENCE: 02307O-081110US
;  CURRENT APPLICATION NUMBER: US/09/191,099
;  CURRENT FILING DATE: 1998-11-12
;  EARLIER APPLICATION NUMBER: US 08/943,571
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 1400
;   TYPE: DNA
;   ORGANISM: Treponema sp.
;   FEATURE:
;   OTHER INFORMATION: 16S rRNA foe spirochete 9-3379
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (21)
;   OTHER INFORMATION: n = unknown
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (24)
;   OTHER INFORMATION: n = unknown
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (26)
;   OTHER INFORMATION: n = unknown
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (122)
;   OTHER INFORMATION: n = unknown
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (173)..(174)
;   OTHER INFORMATION: n = unknown
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (182)
;   OTHER INFORMATION: n = unknown
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (606)
;   OTHER INFORMATION: n = unknown
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (980)
;   OTHER INFORMATION: n = unknown
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (1293)
;   OTHER INFORMATION: n = unknown
US-09-191-099-7

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 1400;
  Best Local Similarity   85.0%;  Pred. No. 40;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
              |||||| ||||| | |||||
Db       1097 GCTTGAGGACTCTGGCGGAA 1116


RESULT 15
US-09-191-099-8
; Sequence 8, Application US/09191099
; Patent No. 6096323
; GENERAL INFORMATION:
;  APPLICANT: Walker, Richard L.
;  APPLICANT: Read, Deryck H.
;  APPLICANT: Hird, David W.
;  APPLICANT: LeFebvre, Rance B.
;  APPLICANT: Berry, Steven L.
;  APPLICANT: Cullor, James S.
;  APPLICANT: Lefler, Hank M.
;  APPLICANT: The Regents of the University of California
;  TITLE OF INVENTION: Vaccine Against Papillomatous Digital Dermatitis (PDD)
;  FILE REFERENCE: 02307O-081110US
;  CURRENT APPLICATION NUMBER: US/09/191,099
;  CURRENT FILING DATE: 1998-11-12
;  EARLIER APPLICATION NUMBER: US 08/943,571
;  EARLIER FILING DATE: 1997-10-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 1446
;   TYPE: DNA
;   ORGANISM: Treponema sp.
US-09-191-099-8

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 1446;
  Best Local Similarity   85.0%;  Pred. No. 41;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GCTTGATGACTCAGCCGGAA 20
```

```
                 |||||| ||||| | |||||
Db       1112 GCTTGAGGACTCTGGCGGAA 1131
```

Search completed: June 26, 2003, 16:20:46
Job time : 40.2888 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 11:19:15 ; Search time 1529.13 Seconds
(without alignments)
211.826 Million cell updates/sec

Title: US-09-355-254F-12
Perfect score: 20
Sequence: 1 tcgatcggggcggggcgagc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	95.0	311	10	BE654158	BE654158 UI-M-AN1-
c	2	19	95.0	429	9	AT005447	AT005447 AT005447
c	3	17.4	87.0	375	10	AW122220	AW122220 UI-M-BH2.
c	4	17.4	87.0	379	9	AI846527	AI846527 UI-M-AN1-
	5	17.4	87.0	406	9	AA615769	AA615769 vo72f08.r
	6	17.4	87.0	423	13	BI220387	BI220387 602935686
	7	17.4	87.0	437	10	BB859799	BB859799 BB859799
	8	17.4	87.0	448	9	AI194235	AI194235 ue52e11.r
	9	17.4	87.0	493	10	BE650594	BE650594 UI-M-BH2.
	10	17.4	87.0	530	12	BF015640	BF015640 uy27b02.y
	11	17.4	87.0	679	13	BI111370	BI111370 602899252
	12	17.4	87.0	709	14	BQ445767	BQ445767 UI-M-ER0-
	13	17.4	87.0	735	14	BQ444735	BQ444735 UI-M-ER0-
	14	17.4	87.0	741	14	BQ445216	BQ445216 UI-M-ER0-
	15	17.4	87.0	764	13	BI146925	BI146925 602911551
	16	17.4	87.0	774	12	BG172820	BG172820 602337255
	17	17.4	87.0	776	14	BQ746273	BQ746273 UI-M-ER0-
	18	17.4	87.0	791	13	BI100531	BI100531 602886667
	19	17.4	87.0	795	12	BG288481	BG288481 602383536
	20	17.4	87.0	795	13	BI332398	BI332398 602983592
	21	17.4	87.0	807	13	BG966084	BG966084 602827404
	22	17.4	87.0	827	13	BI648699	BI648699 603277012
	23	17.4	87.0	832	13	BI144709	BI144709 602909995
	24	17.4	87.0	843	13	BI110250	BI110250 602900674
	25	17.4	87.0	864	12	BF134731	BF134731 601779469
c	26	17.4	87.0	865	13	BI949499	BI949499 HVSME1001
	27	17.4	87.0	883	13	BI328998	BI328998 602980378
	28	17.4	87.0	907	14	BQ930419	BQ930419 AGENCOURT
	29	17.4	87.0	918	12	BF134993	BF134993 601779912
	30	17.4	87.0	919	12	BG244020	BG244020 602358076
	31	17.4	87.0	919	12	BF383950	BF383950 602045373
	32	17.4	87.0	934	13	BI412920	BI412920 602987301
	33	17.4	87.0	938	12	BF384726	BF384726 602046215
	34	17.4	87.0	976	11	AK012729	AK012729 Mus muscu
	35	17.4	87.0	994	12	BF540625	BF540625 602052711
	36	17.4	87.0	997	12	BF580130	BF580130 602099139
	37	17.4	87.0	1073	11	AK020514	AK020514 Mus muscu
	38	17.4	87.0	1081	13	BI852239	BI852239 603378441
	39	17.4	87.0	1083	11	BC012031	BC012031 Mus muscu
	40	17.4	87.0	1101	11	AK010792	AK010792 Mus muscu
	41	17.4	87.0	1114	11	AK020515	AK020515 Mus muscu
	42	17.4	87.0	1115	11	AK019289	AK019289 Mus muscu
	43	17.4	87.0	1254	13	BI554612	BI554612 603235957
	44	17	85.0	508	10	BE232703	BE232703 137723 MA
	45	17	85.0	525	13	BI360501	BI360501 387458 MA

ALIGNMENTS

RESULT 1
BE654158
LOCUS BE654158 311 bp mRNA linear EST 06-SEP-2000
DEFINITION UI-M-AN1-aff-f-04-0-UI.r2 NIH_BMAP_MBG_N Mus musculus cDNA clone UI-M-AN1-aff-f-04-0-UI 5', mRNA sequence.
ACCESSION BE654158
VERSION BE654158.1 GI:9980071
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 311)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose

availability will be considered under appropriate and limited collaborative arrangements The following repetitive elements were found in this cDNA sequence: 119-174, >(CA)n#Simple_repeat
Seq primer: M13 Reverse.

FEATURES Location/Qualifiers
source 1. .311
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-AN1-aff-f-04-0-UI"
/clone_lib="NIH_BMAP_MBG_N"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_MBG_N library is a normalized library constructed from mouse basal ganglia. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6: 791-806 , 1996. Tissue provided by Ms. Annie Novakovich, Zivic-Miller Laboratories."

BASE COUNT 62 a 62 c 116 g 71 t
ORIGIN

Query Match 95.0%; Score 19; DB 10; Length 311;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CGATCGGGGCGGGGCGAGC 20
||||||||||||||||||||
Db 217 CGATCGGGGCGGGGCGAGC 235

RESULT 2
AT005447/c
LOCUS AT005447 429 bp mRNA linear EST 25-MAR-2002
DEFINITION AT005447 POMFBO1 Pleurotus ostreatus cDNA clone MFB34-F01, mRNA sequence.
ACCESSION AT005447
VERSION AT005447.1 GI:13420306
KEYWORDS EST.
SOURCE oyster mushroom.
ORGANISM Pleurotus ostreatus
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Pleurotaceae; Pleurotus.
REFERENCE 1 (bases 1 to 429)
AUTHORS Lee,S.H., Kim,B.G., Kim.K.J., Lee,J.S., Yun,D.W., Hahn,J.H., Kim ,G.H., Lee,K.H., Suh,D.S., Kwon,S.T., Lee,C.S. and Yoo,Y.B.
TITLE Comparative Analysis of Sequences Expressed during the Liquid-Cultured Mycelia and Fruit Body Stages of Pleurotus ostreatus
JOURNAL Fungal Genet. Biol. 35 (2), 115-134 (2002)
MEDLINE 21838665
COMMENT Contact: Beom-Gi Kim
Division of applied microbiology
Institute of Agricultural Science and Technology(NIAST)
249 Seodundong Kweonseonku, Suwon 441707, Korea
Tel: 82-331-290-0347
Fax: 82-331-290-0399
Email: bgkimyes@rda.go.kr
Submitted through BRIC(Biological Research Information Center) of Korea
URL: http://bric.postech.ac.kr/
GeneNuri No. KS105130.

FEATURES Location/Qualifiers
source 1. .429
/organism="Pleurotus ostreatus"
/cultivar="ASI 2029"
/db_xref="taxon:5322"
/clone="MFB34-F01"
/clone_lib="POMFBO1"
/dev_stage="mature fruiting body"
/lab_host="E.coli"
/note="Vector: lambda TripEX2; Site_1: SfiIA; Site_2: SfiIBI; average insert size:1500 bp; initial pfu:5 * 10 7; Isolation of total RNA from the mature fruiting body cultivated in poplar tree sawdust bottle"

BASE COUNT 101 a 135 c 92 g 101 t
ORIGIN

Query Match 95.0%; Score 19; DB 9; Length 429;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGATCGGGGCGGGGCGAG 19
|||||||||||||||||||
Db 119 TCGATCGGGGCGGGGCGAG 101

RESULT 3
AW122220/c
LOCUS AW122220 375 bp mRNA linear EST 22-OCT-1999
DEFINITION UI-M-BH2.2-aov-d-07-0-UI.s1 NIH_BMAP_M_S3.2 Mus musculus cDNA clone UI-M-BH2.2-aov-d-07-0-UI 3', mRNA sequence.
ACCESSION AW122220
VERSION AW122220.1 GI:6097683
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 375)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.
The following repetitive elements were found in this cDNA sequence: 192-247, >(CA)n#Simple_repeat
Seq primer: M13 Forward
POLYA=No.

FEATURES Location/Qualifiers
source 1. .375
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH2.2-aov-d-07-0-UI"
/clone_lib="NIH_BMAP_M_S3.2"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_M_S3.2 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hipoccampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S3.2, NIH_BMAP_M_S2, NIH_BMAP_M_S1.

```
                The subtracted library (NIH_BMAP_M_S3.2) was constructed
                as follows: PCRamplified cDNA inserts from NIH_BMAP_M_S2
                clones from which 3' ESTs had been derived was used as a
                driver in a hybridization with the NIH_BMAP_M_S2 library
                in the form of single-stranded circles. The remaining
                single-stranded circles (subtracted library) was purified
                by hydroxyapatite column chromatography, converted to
                double-stranded circles and electroporated into DH10B
                bacteria (LifeTechnologies) to generate the
                NIH_BMAP_M_S3.2 library. This procedure has been
                previously described (Bonaldo, Lennon and Soares, Genome
                Research 6:791-806, 1996)
                TAG_LIB=NIH_BMAP_M_S3.2
                TAG_TISSUE=cerebellum
                TAG_SEQ=GACTC"
BASE COUNT       87 a    136 c     77 g     75 t
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 10;  Length 375;
  Best Local Similarity   94.7%;  Pred. No. 3.9e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db        100 CGAGCGGGGCGGGGCGAGC 82


RESULT 4
AI846527/c
LOCUS       AI846527                379 bp    mRNA    linear   EST 15-JUL-1999
DEFINITION  UI-M-AN1-aff-f-04-0-UI.s1 NIH_BMAP_MBG_N Mus musculus cDNA clone
            UI-M-AN1-aff-f-04-0-UI 3', mRNA sequence.
ACCESSION   AI846527
VERSION     AI846527.1  GI:5490433
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 379)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
            available by the means that is soon to be determined. When NIH
            determines the means for distribution of the BMAP cDNA clones, this
            record will be updated accordingly when that means is determined.
            The following repetitive elements were found in this cDNA sequence:
            192-247, >(CA)n#Simple_repeat
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .379
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-AN1-aff-f-04-0-UI"
                     /clone_lib="NIH_BMAP_MBG_N"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_MBG_N library is a normalized library constructed
                     from mouse basal ganglia. The tag is a string of 5
                     nucleotides present between the Not I site and the
                     oligo-dT track. The library was constructed as described
                     by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
                     , 1996. Tissue provided by Ms. Annie Novakovich,
                     Zivic-Miller Laboratories.
                     TAG_LIB=NIH_BMAP_MBG_N
                     TAG_TISSUE=hippocampus
                     TAG_SEQ=TTCGA"
BASE COUNT       89 a    136 c     77 g     77 t
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 379;
  Best Local Similarity   94.7%;  Pred. No. 3.9e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db        100 CGAGCGGGGCGGGGCGAGC 82


RESULT 5
AA615769
LOCUS       AA615769                406 bp    mRNA    linear   EST 07-OCT-1997
DEFINITION  vo72f08.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone
            IMAGE:1064679 5' similar to WP:C16C10.7 CE01498 ZINC FINGER PROTEIN
            ;, mRNA sequence.
ACCESSION   AA615769
VERSION     AA615769.1  GI:2502997
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:587039
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 398.
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Mus musculus"
                     /strain="C3H"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1064679"
                     /clone_lib="Barstead mouse myotubes MPLRB5"
                     /cell_line="C2C12"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [AATTCGGATCCTTG], digested with Not I and cloned into the
                     Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead. The C2C12 cell line
                     (available from ATCC, catalog # CRL-1772) differentiates
```

rapidly, forming contractile myotubes and producing characteristic muscle proteins. "
BASE COUNT 81 a 99 c 142 g 83 t 1 others
ORIGIN

Query Match 87.0%; Score 17.4; DB 9; Length 406;
Best Local Similarity 94.7%; Pred. No. 3.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGATCGGGGCGGGGCGAGC 20
||| ||||||||||||||||
Db 158 CGAGCGGGGCGGGGCGAGC 176

RESULT 6
BI220387
LOCUS BI220387 423 bp mRNA linear EST 11-JUL-2001
DEFINITION 602935686F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5098916 5', mRNA sequence.
ACCESSION BI220387
VERSION BI220387.1 GI:14673831
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 423)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11237 row: l column: 21
High quality sequence stop: 419.
FEATURES Location/Qualifiers
source 1. .423
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5098916"
/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 88 a 104 c 155 g 76 t
ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 423;
Best Local Similarity 94.7%; Pred. No. 3.9e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGATCGGGGCGGGGCGAGC 20
||| ||||||||||||||||
Db 138 CGAGCGGGGCGGGGCGAGC 156

RESULT 7
BB859799
LOCUS BB859799 437 bp mRNA linear EST 26-NOV-2001
DEFINITION BB859799 RIKEN full-length enriched, kidney CCL-142 RAG cDNA Mus musculus cDNA clone G430010A01 5', mRNA sequence.
ACCESSION BB859799
VERSION BB859799.1 GI:17101253
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 437)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES Location/Qualifiers
source 1. .437
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="G430010A01"
/clone_lib="RIKEN full-length enriched, kidney CCL-142 RAG cDNA"
/tissue_type="kidney"
/cell_line="CCL-142 RAG"
/note="pooled cell lines ; (cell_line=CRL-1751 WEHI 164), (cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3), (cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA), (cell_line=RCB-0559 K-1 . F1), (cell_line=RCB-1283 B16 melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231), (cell_type=Leydig cells, cell_line=CRL-2065 MLTC-1), (cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE), (tissue_type=bladder, cell_line=RCB-0544 MBT-2), (tissue_type=bone marrow, cell_type=stroma cell, cell_line= CRL-2028 SR-4987), (tissue_type=colon, cell_line=RCB-0549 Cle-H3), (tissue_type=kidney, cell_line=CCL-142 RAG), (tissue_type=submandibular gland, cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C, cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3), (strain=C3H, tissue_type=brain, cell_line=CRL-1443 BC3H1)"
BASE COUNT 89 a 103 c 160 g 85 t
ORIGIN

```
  Query Match            87.0%;  Score 17.4;  DB 10;  Length 437;
  Best Local Similarity  94.7%;  Pred. No. 3.9e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db        168 CGAGCGGGGCGGGGCGAGC 186


RESULT 8
AI194235
LOCUS       AI194235                448 bp    mRNA    linear   EST 13-OCT-1998
DEFINITION  ue52e11.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:1494764 5' similar to TR:O35445 O35445 HYPOTHETICAL 19.8 KD
            PROTEIN. ;, mRNA sequence.
ACCESSION   AI194235
VERSION     AI194235.1  GI:3745442
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:932368
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1494764"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       90 a    110 c    161 g     87 t
ORIGIN

  Query Match            87.0%;  Score 17.4;  DB 9;  Length 448;
  Best Local Similarity  94.7%;  Pred. No. 3.9e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db        143 CGAGCGGGGCGGGGCGAGC 161


RESULT 9
BE650594
LOCUS       BE650594                493 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  UI-M-BH2.2-aov-d-01-0-UI.r1 NIH_BMAP_M_S3.2 Mus musculus cDNA clone
            UI-M-BH2.2-aov-d-01-0-UI 5', mRNA sequence.
ACCESSION   BE650594
VERSION     BE650594.1  GI:9976418
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 493)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
            should be noted that Bento Soares is generating a small number of
            additional specialized non-redundant arrays of BMAP cDNAs whose
            availability will be considered under appropriate and limited
            collaborative arrangements The following repetitive elements were
            found in this cDNA sequence: 118-173, >(CA)n#Simple_repeat
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .493
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH2.2-aov-d-01-0-UI"
                     /clone_lib="NIH_BMAP_M_S3.2"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_M_S3.2 library is a subtracted library of a
                     series,  ultimately derived from a mixture of individually
                     tagged normalized  libraries from ten regions of the mouse
                     brain (cerebellum, brain  stems, olfactory bulbs,
                     hypothalamus, cortex, amygdala, basal  ganglia, pineal
                     gland, striatum, hipoccampus) after a series of
                     subtractions to reduce the representation of cDNAs from
                     which ESTs  had already been generated.  The following
                     serially subtracted  libraries were generated in this
                     process: NIH_BMAP_M_S3.2,  NIH_BMAP_M_S2, NIH_BMAP_M_S1.
                     The subtracted library (NIH_BMAP_M_S3.2)  was constructed
                     as follows: PCRamplified cDNA inserts from NIH_BMAP_M_S2
                     clones from which 3' ESTs had been derived was used  as a
                     driver in a hybridization with the NIH_BMAP_M_S2 library
                     in the  form of single-stranded circles.  The remaining
                     single-stranded  circles (subtracted library) was purified
                     by hydroxyapatite column  chromatography, converted to
                     double-stranded circles and electroporated  into DH10B
                     bacteria (LifeTechnologies) to generate the
                     NIH_BMAP_M_S3.2 library. This procedure has been
                     previously described (Bonaldo,  Lennon and Soares, Genome
                     Research 6:791-806, 1996)"
BASE COUNT       99 a    112 c    158 g    124 t
ORIGIN

  Query Match            87.0%;  Score 17.4;  DB 10;  Length 493;
  Best Local Similarity  94.7%;  Pred. No. 3.8e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db        265 CGAGCGGGGCGGGGCGAGC 283
```

```
RESULT 10
BF015640
LOCUS       BF015640                530 bp    mRNA    linear   EST 29-DEC-2000
DEFINITION  uy27b02.y1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:3660747 5'
            similar to TR:O35445 O35445 HYPOTHETICAL 19.8 KD PROTEIN. ;, mRNA
            sequence.
ACCESSION   BF015640
VERSION     BF015640.1  GI:10746972
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 530)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml

            MGI:1421515
            Seq primer: -40RP from Gibco
            High quality sequence stop: 457.
FEATURES             Location/Qualifiers
     source          1. .530
                     /organism="Mus musculus"
                     /strain="CZECH II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3660747"
                     /clone_lib="NCI_CGAP_Lu30"
                     /tissue_type="tumor, metastatic to mammary"
                     /lab_host="DH10B"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; transgenic model WNT-1, expression driven by
                     MMTV-LTR enhancer; Cloned unidirectionally.  Primer: Oligo
                     dT.   Library constructed by Life Technologies.
                     Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      101 a    135 c    179 g    115 t
ORIGIN

  Query Match            87.0%;  Score 17.4;  DB 12;  Length 530;
  Best Local Similarity   94.7%;  Pred. No. 3.8e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db        144 CGAGCGGGGCGGGGCGAGC 162


RESULT 11
BI111370
LOCUS       BI111370                679 bp    mRNA    linear   EST 26-JUN-2001
DEFINITION  602899252F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5028924 5',
            mRNA sequence.
ACCESSION   BI111370
VERSION     BI111370.1  GI:14562271
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 679)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11081  row: h  column: 13
            High quality sequence stop: 666.
FEATURES             Location/Qualifiers
     source          1. .679
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5028924"
                     /clone_lib="NCI_CGAP_Mam5"
                     /tissue_type="tumor, gross tissue"
                     /dev_stage="7 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigators
                     providing samples: Lothar Hennighausen/Robin Humphreys,
                     NIH"
BASE COUNT      126 a    190 c    216 g    146 t      1 others
ORIGIN

  Query Match            87.0%;  Score 17.4;  DB 13;  Length 679;
  Best Local Similarity   94.7%;  Pred. No. 3.8e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db        129 CGAGCGGGGCGGGGCGAGC 147


RESULT 12
BQ445767
LOCUS       BQ445767                709 bp    mRNA    linear   EST 29-MAY-2002
DEFINITION  UI-M-ER0-bxm-g-18-0-UI.r2 NIH_BMAP_ER0 Mus musculus cDNA clone
            IMAGE:5710121 5', mRNA sequence.
ACCESSION   BQ445767
VERSION     BQ445767.1  GI:21248879
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 709)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, Univeristy of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .709
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
```

```
                     /clone="IMAGE:5710121"
                     /clone_lib="NIH_BMAP_ER0"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 15.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured mRNa was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was size selected according to mRNA size fraction,
                     ligated with EcoR I adaptor, digested with Not I, and then
                     cloned directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     , is GTGCGTGGAA. This library was created for the
                     University of Iowa Mouse Brain Molecular Anatomy Project
                     (BMAP): 'Gene Discovery in the Developing Mouse Nervous
                     System', supported by National Instututes of Mental Health
                     (NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT      131 a    201 c    211 g    164 t      2 others
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 14;  Length 709;
  Best Local Similarity   94.7%;  Pred. No. 3.8e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db         90 CGAGCGGGGCGGGGCGAGC 108


RESULT 13
BQ444735
LOCUS       BQ444735                735 bp    mRNA    linear   EST 29-MAY-2002
DEFINITION  UI-M-ER0-bxm-h-21-0-UI.r1 NIH_BMAP_ER0 Mus musculus cDNA clone
            IMAGE:5710148 5', mRNA sequence.
ACCESSION   BQ444735
VERSION     BQ444735.1  GI:21247847
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 735)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, Univeristy of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .735
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5710148"
                     /clone_lib="NIH_BMAP_ER0"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 15.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured mRNa was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was size selected according to mRNA size fraction,
                     ligated with EcoR I adaptor, digested with Not I, and then
                     cloned directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     , is GTGCGTGGAA. This library was created for the
                     University of Iowa Mouse Brain Molecular Anatomy Project
                     (BMAP): 'Gene Discovery in the Developing Mouse Nervous
                     System', supported by National Instututes of Mental Health
                     (NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT      135 a    205 c    225 g    168 t      2 others
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 14;  Length 735;
  Best Local Similarity   94.7%;  Pred. No. 3.8e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CGATCGGGGCGGGGCGAGC 20
              ||| ||||||||||||||||
Db         90 CGAGCGGGGCGGGGCGAGC 108


RESULT 14
BQ445216
LOCUS       BQ445216                741 bp    mRNA    linear   EST 29-MAY-2002
DEFINITION  UI-M-ER0-bxp-d-17-0-UI.r1 NIH_BMAP_ER0 Mus musculus cDNA clone
            IMAGE:5711200 5', mRNA sequence.
ACCESSION   BQ445216
VERSION     BQ445216.1  GI:21248328
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 741)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin, Univeristy of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
             This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source          1. .741
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5711200"
                     /clone_lib="NIH_BMAP_ER0"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 15.5 dpc"
                     /lab_host="DH10B (T1 phage resistant)"
                     /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
                     Site_2: Not I; The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured mRNa was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was size selected according to mRNA size fraction,
                     ligated with EcoR I adaptor, digested with Not I, and then
                     cloned directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     , is GTGCGTGGAA. This library was created for the
```

University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Instututes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT 133 a 202 c 224 g 179 t 3 others
ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 741;
Best Local Similarity 94.7%; Pred. No. 3.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGATCGGGGCGGGGCGAGC 20
||| ||||||||||||||||
Db 129 CGAGCGGGGCGGGGCGAGC 147

RESULT 15
BI146925
LOCUS BI146925 764 bp mRNA linear EST 05-JUL-2001
DEFINITION 602911551F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5052660 5', mRNA sequence.
ACCESSION BI146925
VERSION BI146925.1 GI:14606926
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 764)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11143 row: e column: 13
High quality sequence stop: 741.
FEATURES Location/Qualifiers
source 1. .764
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5052660"
/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 138 a 211 c 235 g 180 t
ORIGIN

Query Match 87.0%; Score 17.4; DB 13; Length 764;
Best Local Similarity 94.7%; Pred. No. 3.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CGATCGGGGCGGGGCGAGC 20
||| ||||||||||||||||
Db 115 CGAGCGGGGCGGGGCGAGC 133

Search completed: June 26, 2003, 22:12:31
Job time : 1533.13 secs

GenCore version 5.1.6,
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 10:58:25 ; Search time 30.2174 Seconds
(without alignments)
202.980 Million cell updates/sec

Title: US-09-355-254F-13
Perfect score: 20
Sequence: 1 tgcagattgcgcaatctgca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	4	US-09-021-247-7	Sequence 7, Appli
c	2	20	100.0	20	4	US-09-021-247-7	Sequence 7, Appli
	3	20	100.0	20	4	US-09-425-798-15	Sequence 15, Appl
c	4	20	100.0	20	4	US-09-425-798-15	Sequence 15, Appl
	5	14.8	74.0	653	4	US-08-998-416-1111	Sequence 1111, Ap
c	6	14.8	74.0	653	4	US-08-998-416-1111	Sequence 1111, Ap
	7	14.4	72.0	49	1	US-07-601-094-5	Sequence 5, Appli
c	8	14.4	72.0	49	1	US-07-601-094-5	Sequence 5, Appli
	9	14.4	72.0	49	1	US-08-012-735-5	Sequence 5, Appli
c	10	14.4	72.0	49	1	US-08-012-735-5	Sequence 5, Appli
	11	14.2	71.0	522	2	US-08-767-026-1	Sequence 1, Appli
c	12	14.2	71.0	522	2	US-08-767-026-1	Sequence 1, Appli
	13	14.2	71.0	857	4	US-08-998-416-558	Sequence 558, App
c	14	14.2	71.0	857	4	US-08-998-416-558	Sequence 558, App
	15	14.2	71.0	1101	4	US-09-210-843-1	Sequence 1, Appli
c	16	14.2	71.0	1101	4	US-09-210-843-1	Sequence 1, Appli
	17	14.2	71.0	1800	1	US-08-366-783-1	Sequence 1, Appli
c	18	14.2	71.0	1800	1	US-08-366-783-1	Sequence 1, Appli
	19	14.2	71.0	1800	1	US-08-313-098A-1	Sequence 1, Appli
c	20	14.2	71.0	1800	1	US-08-313-098A-1	Sequence 1, Appli
	21	14.2	71.0	1800	2	US-08-846-021A-1	Sequence 1, Appli
c	22	14.2	71.0	1800	2	US-08-846-021A-1	Sequence 1, Appli
	23	14.2	71.0	1897	6	RE34606-5	Patent No. RE34,60
c	24	14.2	71.0	1897	6	RE34606-5	Patent No. RE34,60
	25	14.2	71.0	2115	2	US-08-767-026-3	Sequence 3, Appli
c	26	14.2	71.0	2115	2	US-08-767-026-3	Sequence 3, Appli
	27	14.2	71.0	2133	4	US-09-488-744A-3	Sequence 3, Appli
c	28	14.2	71.0	2133	4	US-09-488-744A-3	Sequence 3, Appli
	29	14.2	71.0	2733	2	US-08-846-021A-6	Sequence 6, Appli
c	30	14.2	71.0	2733	2	US-08-846-021A-6	Sequence 6, Appli
	31	14.2	71.0	3183	1	US-08-849-212-3	Sequence 3, Appli
c	32	14.2	71.0	3183	1	US-08-849-212-3	Sequence 3, Appli
	33	14.2	71.0	3546	4	US-09-118-442-14	Sequence 14, Appl
c	34	14.2	71.0	3546	4	US-09-118-442-14	Sequence 14, Appl
	35	14.2	71.0	3546	4	US-09-118-442-15	Sequence 15, Appl
c	36	14.2	71.0	3546	4	US-09-118-442-15	Sequence 15, Appl
	37	14.2	71.0	3546	4	US-09-677-064-14	Sequence 14, Appl
c	38	14.2	71.0	3546	4	US-09-677-064-14	Sequence 14, Appl
	39	14.2	71.0	3546	4	US-09-677-064-15	Sequence 15, Appl
c	40	14.2	71.0	3546	4	US-09-677-064-15	Sequence 15, Appl
	41	14.2	71.0	9578	4	US-08-961-527-127	Sequence 127, App
c	42	14.2	71.0	9578	4	US-08-961-527-127	Sequence 127, App
	43	14	70.0	960	2	US-08-245-511-3	Sequence 3, Appli
c	44	14	70.0	960	2	US-08-245-511-3	Sequence 3, Appli
	45	14	70.0	960	2	US-08-600-993A-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-09-021-247-7
; Sequence 7, Application US/09021247
; Patent No. 6225444
;  GENERAL INFORMATION:
;    APPLICANT:  Shashoua, Victor E.
;    TITLE OF INVENTION:  NEUROPROTECTIVE PEPTIDES AND USES THEREOF
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/021,247
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Van Amsterdam, John R.
;      REGISTRATION NUMBER:  40,212
;      REFERENCE/DOCKET NUMBER:  N0260/7023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  oligonucleotide
;    HYPOTHETICAL:  NO
US-09-021-247-7

  Query Match           100.0%;  Score 20;  DB 4;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 0.083;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              ||||||||||||||||||||
Db          1 TGCAGATTGCGCAATCTGCA 20
```

```
RESULT 2
US-09-021-247-7/c
; Sequence 7, Application US/09021247
; Patent No. 6225444
;  GENERAL INFORMATION:
;    APPLICANT:  Shashoua, Victor E.
;    TITLE OF INVENTION:  NEUROPROTECTIVE PEPTIDES AND USES THEREOF
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/021,247
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Van Amsterdam, John R.
;      REGISTRATION NUMBER:  40,212
;      REFERENCE/DOCKET NUMBER:  N0260/7023
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 nucleotides
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  oligonucleotide
;    HYPOTHETICAL:  NO
US-09-021-247-7

  Query Match             100.0%;  Score 20;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.083;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              ||||||||||||||||||||
Db         20 TGCAGATTGCGCAATCTGCA 1


RESULT 3
US-09-425-798-15
; Sequence 15, Application US/09425798A
; Patent No. 6423493
; GENERAL INFORMATION:
;  APPLICANT:  Gorenstein Dr., David G.
;  APPLICANT:  King Dr., David J.
;  APPLICANT:  Ventura, Daniel A.
;  APPLICANT:  Brasier Dr., Allan R.
;  TITLE OF INVENTION:  Combinatorial Selection of Phosphothionate
;  TITLE OF INVENTION:  Oligonucleotide Aptamers
;  FILE REFERENCE:  122144-1005
;  CURRENT APPLICATION NUMBER: US/09/425,798A
;  CURRENT FILING DATE:  1999-10-25
;  PRIOR APPLICATION NUMBER:  60/105,600
;  PRIOR FILING DATE:  1999-10-26
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: aptamer
US-09-425-798-15

  Query Match             100.0%;  Score 20;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.083;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              ||||||||||||||||||||
Db          1 TGCAGATTGCGCAATCTGCA 20


RESULT 4
US-09-425-798-15/c
; Sequence 15, Application US/09425798A
; Patent No. 6423493
; GENERAL INFORMATION:
;  APPLICANT:  Gorenstein Dr., David G.
;  APPLICANT:  King Dr., David J.
;  APPLICANT:  Ventura, Daniel A.
;  APPLICANT:  Brasier Dr., Allan R.
;  TITLE OF INVENTION:  Combinatorial Selection of Phosphothionate
;  TITLE OF INVENTION:  Oligonucleotide Aptamers
;  FILE REFERENCE:  122144-1005
;  CURRENT APPLICATION NUMBER: US/09/425,798A
;  CURRENT FILING DATE:  1999-10-25
;  PRIOR APPLICATION NUMBER:  60/105,600
;  PRIOR FILING DATE:  1999-10-26
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: aptamer
US-09-425-798-15

  Query Match             100.0%;  Score 20;  DB 4;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.083;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              ||||||||||||||||||||
Db         20 TGCAGATTGCGCAATCTGCA 1


RESULT 5
US-08-998-416-1111
; Sequence 1111, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  1111:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  653 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1677RP
US-08-998-416-1111

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 653;
  Best Local Similarity   88.9%;  Pred. No. 73;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCAGATTGCGCAATCTGC 19
              |||||||| ||||||| |
Db        126 GCAGATTGAGCAATCTTC 143


RESULT 6
US-08-998-416-1111/c
; Sequence 1111, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinnè
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  1111:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  653 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1677RP
US-08-998-416-1111

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 653;
  Best Local Similarity   88.9%;  Pred. No. 73;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCAGATTGCGCAATCTGC 19
              | ||||||| ||||||||
Db        143 GAAGATTGCTCAATCTGC 126


RESULT 7
US-07-601-094-5
; Sequence 5, Application US/07601094
; Patent No. 5215892
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Hirano, Toshio
;    APPLICANT:  Akira, Shizuo
;    APPLICANT:  Isshiki, Hiroshi
;    APPLICANT:  Tanabe, Osamu
;    APPLICANT:  Kinoshita, Shigemi
;    APPLICANT:  Shimamoto, Takuya
;    TITLE OF INVENTION:  C/EBP2 Gene and Recombinant
;    TITLE OF INVENTION:  C/EBP2
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak &
;      ADDRESSEE:  Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/601,094
;      FILING DATE:  19901022
;      CLASSIFICATION:  435
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-07-601-094-5

  Query Match             72.0%;  Score 14.4;  DB 1;  Length 49;
  Best Local Similarity   93.8%;  Pred. No. 80;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 CAGATTGCGCAATCTG 18
```

```
                ||||||||| |||||||
Db          18 CAGATTGCACAATCTG 33


RESULT 8
US-07-601-094-5/c
; Sequence 5, Application US/07601094
; Patent No. 5215892
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Hirano, Toshio
;    APPLICANT:  Akira, Shizuo
;    APPLICANT:  Isshiki, Hiroshi
;    APPLICANT:  Tanabe, Osamu
;    APPLICANT:  Kinoshita, Shigemi
;    APPLICANT:  Shimamoto, Takuya
;    TITLE OF INVENTION:  C/EBP2 Gene and Recombinant
;    TITLE OF INVENTION:  C/EBP2
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak &
;      ADDRESSEE:  Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/601,094
;      FILING DATE:  19901022
;      CLASSIFICATION:  435
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-07-601-094-5

  Query Match             72.0%;  Score 14.4;  DB 1;  Length 49;
  Best Local Similarity   93.8%;  Pred. No. 80;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 CAGATTGCGCAATCTG 18
              ||||||| ||||||||
Db         33 CAGATTGTGCAATCTG 18


RESULT 9
US-08-012-735-5
; Sequence 5, Application US/08012735
; Patent No. 5360894
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Hirano, Toshio
;    APPLICANT:  Akira, Shizuo
;    APPLICANT:  Isshiki, Hiroshi
;    APPLICANT:  Tanabe, Osamu
;    APPLICANT:  Kinoshita, Shigemi
;    APPLICANT:  Shimamoto, Takuya
;    TITLE OF INVENTION:  C/EBP2 Gene and Recombinant
;    TITLE OF INVENTION:  C/EBP2
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak &
;      ADDRESSEE:  Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/012,735
;      FILING DATE:  19930203
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/601,094
;      FILING DATE:  22 OCT 1990
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-012-735-5

  Query Match             72.0%;  Score 14.4;  DB 1;  Length 49;
  Best Local Similarity   93.8%;  Pred. No. 80;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 CAGATTGCGCAATCTG 18
              ||||||| ||||||||
Db         18 CAGATTGCACAATCTG 33


RESULT 10
US-08-012-735-5/c
; Sequence 5, Application US/08012735
; Patent No. 5360894
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Hirano, Toshio
;    APPLICANT:  Akira, Shizuo
;    APPLICANT:  Isshiki, Hiroshi
;    APPLICANT:  Tanabe, Osamu
;    APPLICANT:  Kinoshita, Shigemi
;    APPLICANT:  Shimamoto, Takuya
;    TITLE OF INVENTION:  C/EBP2 Gene and Recombinant
;    TITLE OF INVENTION:  C/EBP2
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak &
;      ADDRESSEE:  Seas
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States
;      ZIP:  20037-3202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/012,735
;      FILING DATE:  19930203
```

```
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/601,094
;      FILING DATE:  22 OCT 1990
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 293-7060
;      TELEFAX:  (202) 293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-012-735-5

  Query Match             72.0%;  Score 14.4;  DB 1;  Length 49;
  Best Local Similarity   93.8%;  Pred. No. 80;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 CAGATTGCGCAATCTG 18
              ||||||| ||||||||
Db         33 CAGATTGTGCAATCTG 18


RESULT 11
US-08-767-026-1
; Sequence 1, Application US/08767026
; Patent No. 5856452
;  GENERAL INFORMATION:
;    APPLICANT:  Moloney, Maurice
;    APPLICANT:  Boothe, Joseph
;    APPLICANT:  van Rooijen, Gijs
;    TITLE OF INVENTION:  Oil Bodies and Associated Proteins as
;    TITLE OF INVENTION:  Affinity Matrices
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BERESKIN & PARR
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/767,026
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  9369-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  522 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Oleosin From Arabidopsis Thaliana
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..522
US-08-767-026-1

  Query Match             71.0%;  Score 14.2;  DB 2;  Length 522;
  Best Local Similarity   84.2%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GCAGATTGCGCAATCTGCA 20
              |||||||||  || |||||
Db        123 GCAGATTGCTAAAGCTGCA 141


RESULT 12
US-08-767-026-1/c
; Sequence 1, Application US/08767026
; Patent No. 5856452
;  GENERAL INFORMATION:
;    APPLICANT:  Moloney, Maurice
;    APPLICANT:  Boothe, Joseph
;    APPLICANT:  van Rooijen, Gijs
;    TITLE OF INVENTION:  Oil Bodies and Associated Proteins as
;    TITLE OF INVENTION:  Affinity Matrices
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BERESKIN & PARR
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/767,026
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gravelle, Micheline
;      REGISTRATION NUMBER:  40,261
;      REFERENCE/DOCKET NUMBER:  9369-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  522 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Oleosin From Arabidopsis Thaliana
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..522
US-08-767-026-1

  Query Match             71.0%;  Score 14.2;  DB 2;  Length 522;
  Best Local Similarity   84.2%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGC 19
              ||||| ||  |||||||||
Db        141 TGCAGCTTTAGCAATCTGC 123


RESULT 13
US-08-998-416-558
; Sequence 558, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
```

```
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  558:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  857 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1387UP
US-08-998-416-558

  Query Match             71.0%;  Score 14.2;  DB 4;  Length 857;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              |||||   ||||||||| ||
Db        699 TGCAGCNGGCGCAATCTACA 718


RESULT 14
US-08-998-416-558/c
; Sequence 558, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  558:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  857 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1387UP
US-08-998-416-558

  Query Match             71.0%;  Score 14.2;  DB 4;  Length 857;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              || |||||||||   |||||
Db        718 TGTAGATTGCGCCNGCTGCA 699


RESULT 15
US-09-210-843-1
; Sequence 1, Application US/09210843
; Patent No. 6288304
; GENERAL INFORMATION:
;  APPLICANT: Moloney, Maurice M.
;  APPLICANT: Habibi, Hamid R.
;  TITLE OF INVENTION: Expression  of Somatotropin in Plant Seeds
;  FILE REFERENCE: 9369-69
;  CURRENT APPLICATION NUMBER: US/09/210,843
;  CURRENT FILING DATE: 1998-12-15
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1101
;   TYPE: DNA
;   ORGANISM: synthetic construct
US-09-210-843-1

  Query Match             71.0%;  Score 14.2;  DB 4;  Length 1101;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GCAGATTGCGCAATCTGCA 20
              |||||||||  || |||||
Db        123 GCAGATTGCTAAAGCTGCA 141


Search completed: June 26, 2003, 16:20:52
Job time : 31.2888 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 11:19:15 ; Search time 1529.13 Seconds
(without alignments)
211.826 Million cell updates/sec

Title: US-09-355-254F-13
Perfect score: 20
Sequence: 1 tgcagattgcgcaatctgca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	16.8	84.0	421	10	AW440244	AW440244 xu42h12.x
c	2	16.8	84.0	421	10	AW440244	AW440244 xu42h12.x
	3	16.8	84.0	456	17	AQ696464	AQ696464 HS_5518_A
c	4	16.8	84.0	456	17	AQ696464	AQ696464 HS_5518_A
	5	16.4	82.0	380	17	BH381085	BH381085 AG-ND-157
c	6	16.4	82.0	380	17	BH381085	BH381085 AG-ND-157
	7	16.4	82.0	478	17	BH372378	BH372378 AG-ND-101
c	8	16.4	82.0	478	17	BH372378	BH372378 AG-ND-101
	9	15.8	79.0	266	17	AZ553458	AZ553458 RPCI-23-2
c	10	15.8	79.0	266	17	AZ553458	AZ553458 RPCI-23-2
	11	15.8	79.0	269	9	AI913095	AI913095 tz88a12.x
c	12	15.8	79.0	269	9	AI913095	AI913095 tz88a12.x
	13	15.8	79.0	301	10	BB071586	BB071586 BB071586
c	14	15.8	79.0	301	10	BB071586	BB071586 BB071586
	15	15.8	79.0	393	14	BU026495	BU026495 QHG17A10.
c	16	15.8	79.0	393	14	BU026495	BU026495 QHG17A10.
	17	15.8	79.0	424	13	BM146790	BM146790 TCAAP1E73
c	18	15.8	79.0	424	13	BM146790	BM146790 TCAAP1E73
	19	15.8	79.0	463	17	AQ176460	AQ176460 HS_3207_A
c	20	15.8	79.0	463	17	AQ176460	AQ176460 HS_3207_A
	21	15.8	79.0	520	10	AW447996	AW447996 BRY_1196
c	22	15.8	79.0	520	10	AW447996	AW447996 BRY_1196
	23	15.8	79.0	520	14	BQ605683	BQ605683 BRY_1196
c	24	15.8	79.0	520	14	BQ605683	BQ605683 BRY_1196
	25	15.8	79.0	633	12	BG568188	BG568188 602587084
c	26	15.8	79.0	633	12	BG568188	BG568188 602587084
	27	15.8	79.0	724	17	AG184160	AG184160 Pan trogl
c	28	15.8	79.0	724	17	AG184160	AG184160 Pan trogl
	29	15.8	79.0	928	14	BQ731879	BQ731879 AGENCOURT
c	30	15.8	79.0	928	14	BQ731879	BQ731879 AGENCOURT
	31	15.8	79.0	1157	11	AK020226	AK020226 Mus muscu
c	32	15.8	79.0	1157	11	AK020226	AK020226 Mus muscu
	33	15.4	77.0	242	12	BF554724	BF554724 UI-R-C0-j
c	34	15.4	77.0	242	12	BF554724	BF554724 UI-R-C0-j
	35	15.4	77.0	367	12	BF004180	BF004180 EST432678
c	36	15.4	77.0	367	12	BF004180	BF004180 EST432678
	37	15.4	77.0	404	9	AI563420	AI563420 EST00544
c	38	15.4	77.0	404	9	AI563420	AI563420 EST00544
	39	15.4	77.0	481	17	B97611	B97611 F15N19TFB I
c	40	15.4	77.0	481	17	B97611	B97611 F15N19TFB I
	41	15.4	77.0	501	10	BE240794	BE240794 EST404843
c	42	15.4	77.0	501	10	BE240794	BE240794 EST404843
	43	15.4	77.0	610	10	AW773826	AW773826 EST332812
c	44	15.4	77.0	610	10	AW773826	AW773826 EST332812
	45	15.4	77.0	698	12	BG853315	BG853315 1024036G0

ALIGNMENTS

RESULT 1
AW440244
LOCUS AW440244 421 bp mRNA linear EST 14-FEB-2000
DEFINITION xu42h12.x1 NCI_CGAP_HN9 Homo sapiens cDNA clone IMAGE:2804423 3' similar to contains L1.t2 L1 repetitive element ;, mRNA sequence.
ACCESSION AW440244
VERSION AW440244.1 GI:6975550
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 421)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D., Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found

```
           Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2804423"
                     /clone_lib="NCI_CGAP_HN9"
                     /tissue_type="normal squamous epithelium from retromolar
                     trigone"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; cDNA made by oligo-dT priming.
                     Non-directionally cloned into the UDG sites of pAMP10.
                     Size-selected on agarose gel, average insert  size 500 bp.
                     Primary library; non-amplified.   cDNA Library
                     Preparation: David B. Krizman, Ph.D (NCI). Reference:
                     Krizman et al. (1996) Cancer Research 56:5380-5383."
BASE COUNT      159 a     88 c     81 g     93 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 10;  Length 421;
  Best Local Similarity   90.0%;  Pred. No. 5.9e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              ||||||||| | ||||||||
Db        125 TGCAGATTGTGAAATCTGCA 144


RESULT 2
AW440244/c
LOCUS       AW440244                421 bp    mRNA    linear   EST 14-FEB-2000
DEFINITION  xu42h12.x1 NCI_CGAP_HN9 Homo sapiens cDNA clone IMAGE:2804423 3'
            similar to contains L1.t2 L1 repetitive element ;, mRNA sequence.
ACCESSION   AW440244
VERSION     AW440244.1  GI:6975550
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Edward Shillitoe Ph.D., Silvio Gutkind Ph.D.,
            Chidchanok Leethanakul D.D.S., Michael Emmert-Buck M.D. Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2804423"
                     /clone_lib="NCI_CGAP_HN9"
                     /tissue_type="normal squamous epithelium from retromolar
                     trigone"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; cDNA made by oligo-dT priming.
                     Non-directionally cloned into the UDG sites of pAMP10.
                     Size-selected on agarose gel, average insert  size 500 bp.
                     Primary library; non-amplified.   cDNA Library
                     Preparation: David B. Krizman, Ph.D (NCI). Reference:
                     Krizman et al. (1996) Cancer Research 56:5380-5383."
BASE COUNT      159 a     88 c     81 g     93 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 10;  Length 421;
  Best Local Similarity   90.0%;  Pred. No. 5.9e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              ||||||||| | ||||||||
Db        144 TGCAGATTTCACAATCTGCA 125


RESULT 3
AQ696464
LOCUS       AQ696464                456 bp    DNA     linear   GSS 06-JUL-1999
DEFINITION  HS_5518_A2_D04_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1094 Col=8 Row=G, DNA sequence.
ACCESSION   AQ696464
VERSION     AQ696464.1  GI:5386712
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1094  row: G  column: 8
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 456.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=1094 Col=8 Row=G"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      142 a     92 c    103 g    118 t      1 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 17;  Length 456;
  Best Local Similarity   90.0%;  Pred. No. 6.1e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              |||||||||| || ||||||
Db        365 TGCAGATTGCACATTCTGCA 384
```

```
RESULT 4
AQ696464/c
LOCUS       AQ696464                456 bp    DNA     linear   GSS 06-JUL-1999
DEFINITION  HS_5518_A2_D04_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1094 Col=8 Row=G, DNA sequence.
ACCESSION   AQ696464
VERSION     AQ696464.1  GI:5386712
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1094  row: G  column: 8
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 456.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=1094 Col=8 Row=G"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      142 a     92 c    103 g    118 t      1 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 17;  Length 456;
  Best Local Similarity   90.0%;  Pred. No. 6.1e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGCA 20
              |||||| || ||||||||||
Db        384 TGCAGAATGTGCAATCTGCA 365


RESULT 5
BH381085
LOCUS       BH381085                380 bp    DNA     linear   GSS 10-DEC-2001
DEFINITION  AG-ND-157L22.TF ND-TAM Anopheles gambiae genomic clone AG-ND-157L22
            , DNA sequence.
ACCESSION   BH381085
VERSION     BH381085.1  GI:17327227
KEYWORDS    GSS.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Anopheles.
REFERENCE   1  (bases 1 to 380)
  AUTHORS   Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
  TITLE     Direct Submission of BAC-end sequences from Anopheles gambiae
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: AG-ND-157L22.TR
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            This clone is from an A. gambiae BAC library (ND-TAM) provided by
            F.H. Collins and sequenced by The Institute for Genomic Research
            (TIGR). The BAC library was generated from A. gambiae PEST strain
            DNA. All DNA was extracted from newly hatched first instar larvae
            to minimize the inclusion of DNA from microorganisms that inhabit
            the gut. The DNA is derived from mixed sexes of larvae. The BAC
            library was constructed at Texas A&M Universty BAC Center
            University, College Station, Texas 77843-2123, USA using a HindIII
            partial digest.
            Seq primer: M13 For
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .380
                     /organism="Anopheles gambiae"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="AG-ND-157L22"
                     /clone_lib="ND-TAM"
                     /note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT       69 a    111 c    119 g     81 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 17;  Length 380;
  Best Local Similarity   94.4%;  Pred. No. 8.9e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 CAGATTGCGCAATCTGCA 20
              |||||| ||||||||||
Db        280 CAGATTCCGCAATCTGCA 297


RESULT 6
BH381085/c
LOCUS       BH381085                380 bp    DNA     linear   GSS 10-DEC-2001
DEFINITION  AG-ND-157L22.TF ND-TAM Anopheles gambiae genomic clone AG-ND-157L22
            , DNA sequence.
ACCESSION   BH381085
VERSION     BH381085.1  GI:17327227
KEYWORDS    GSS.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Anopheles.
REFERENCE   1  (bases 1 to 380)
  AUTHORS   Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
  TITLE     Direct Submission of BAC-end sequences from Anopheles gambiae
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: AG-ND-157L22.TR
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            This clone is from an A. gambiae BAC library (ND-TAM) provided by
            F.H. Collins and sequenced by The Institute for Genomic Research
            (TIGR). The BAC library was generated from A. gambiae PEST strain
            DNA. All DNA was extracted from newly hatched first instar larvae
```

```
          to minimize the inclusion of DNA from microorganisms that inhabit
          the gut. The DNA is derived from mixed sexes of larvae. The BAC
          library was constructed at Texas A&M Universty BAC Center
          University, College Station, Texas 77843-2123, USA using a HindIII
          partial digest.
          Seq primer: M13 For
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .380
                     /organism="Anopheles gambiae"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="AG-ND-157L22"
                     /clone_lib="ND-TAM"
                     /note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT       69 a    111 c    119 g     81 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 17;  Length 380;
  Best Local Similarity   94.4%;  Pred. No. 8.9e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTG 18
              |||||||||||  ||||||
Db        297 TGCAGATTGCGGAATCTG 280


RESULT 7
BH372378
LOCUS       BH372378                478 bp    DNA     linear   GSS 10-DEC-2001
DEFINITION  AG-ND-101M20.TF ND-TAM Anopheles gambiae genomic clone AG-ND-101M20
          , DNA sequence.
ACCESSION   BH372378
VERSION     BH372378.1  GI:17318503
KEYWORDS    GSS.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
          Anopheles.
REFERENCE   1  (bases 1 to 478)
  AUTHORS   Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
  TITLE     Direct Submission of BAC-end sequences from Anopheles gambiae
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: AG-ND-101M20.TR
          Contact: Brendan J Loftus
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0208
          Fax: 301 838 3543
          Email: bjloftus@tigr.org
          This clone is from an A. gambiae BAC library (ND-TAM) provided by
          F.H. Collins and sequenced by The Institute for Genomic Research
          (TIGR). The BAC library was generated from A. gambiae PEST strain
          DNA. All DNA was extracted from newly hatched first instar larvae
          to minimize the inclusion of DNA from microorganisms that inhabit
          the gut. The DNA is derived from mixed sexes of larvae. The BAC
          library was constructed at Texas A&M Universty BAC Center
          University, College Station, Texas 77843-2123, USA using a HindIII
          partial digest.
          Seq primer: M13 For
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .478
                     /organism="Anopheles gambiae"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="AG-ND-101M20"
                     /clone_lib="ND-TAM"
                     /note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT      153 a    109 c    123 g     93 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 17;  Length 478;
  Best Local Similarity   94.4%;  Pred. No. 9.9e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 CAGATTGCGCAATCTGCA 20
              | ||||||||||||||||
Db         17 CCGATTGCGCAATCTGCA 34


RESULT 8
BH372378/c
LOCUS       BH372378                478 bp    DNA     linear   GSS 10-DEC-2001
DEFINITION  AG-ND-101M20.TF ND-TAM Anopheles gambiae genomic clone AG-ND-101M20
          , DNA sequence.
ACCESSION   BH372378
VERSION     BH372378.1  GI:17318503
KEYWORDS    GSS.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
          Anopheles.
REFERENCE   1  (bases 1 to 478)
  AUTHORS   Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
  TITLE     Direct Submission of BAC-end sequences from Anopheles gambiae
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: AG-ND-101M20.TR
          Contact: Brendan J Loftus
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0208
          Fax: 301 838 3543
          Email: bjloftus@tigr.org
          This clone is from an A. gambiae BAC library (ND-TAM) provided by
          F.H. Collins and sequenced by The Institute for Genomic Research
          (TIGR). The BAC library was generated from A. gambiae PEST strain
          DNA. All DNA was extracted from newly hatched first instar larvae
          to minimize the inclusion of DNA from microorganisms that inhabit
          the gut. The DNA is derived from mixed sexes of larvae. The BAC
          library was constructed at Texas A&M Universty BAC Center
          University, College Station, Texas 77843-2123, USA using a HindIII
          partial digest.
          Seq primer: M13 For
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .478
                     /organism="Anopheles gambiae"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="AG-ND-101M20"
                     /clone_lib="ND-TAM"
                     /note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT      153 a    109 c    123 g     93 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 17;  Length 478;
  Best Local Similarity   94.4%;  Pred. No. 9.9e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTG 18
              |||||||||||||||| |
Db         34 TGCAGATTGCGCAATCGG 17


RESULT 9
AZ553458
LOCUS       AZ553458                266 bp    DNA     linear   GSS 20-NOV-2000
DEFINITION  RPCI-23-211A16.TV RPCI-23 Mus musculus genomic clone RPCI-23-211A16
          , DNA sequence.
ACCESSION   AZ553458
VERSION     AZ553458.1  GI:11233049
```

```
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 266)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-211A16.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 211  row: A  column: 16
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .266
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="RPCI-23-211A16"
                      /clone_lib="RPCI-23"
                      /sex="Female"
                      /lab_host="DH10B"
                      /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                      EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                      brain genomic DNA was isolated and  partially digested
                      with a combination of EcoRI and EcoRI Methylase.   Size
                      selected DNA was cloned into the pBACe3.6 vector at the
                      EcoRI  sites.  The ligation products were transformed into
                      DH10B  electrocompetent cells (BRL Life Technologies).  "
BASE COUNT       88 a     49 c     40 g     89 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 17;  Length 266;
  Best Local Similarity   89.5%;  Pred. No. 1.5e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCAGATTGCGCAATCTGCA 20
              |||||| || |||||||||
Db         12 GCAGATAGCACAATCTGCA 30


RESULT 10
AZ553458/c
LOCUS       AZ553458               266 bp    DNA     linear   GSS 20-NOV-2000
DEFINITION  RPCI-23-211A16.TV RPCI-23 Mus musculus genomic clone RPCI-23-211A16
            , DNA sequence.
ACCESSION   AZ553458
VERSION     AZ553458.1  GI:11233049
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 266)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-211A16.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 211  row: A  column: 16
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .266
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="RPCI-23-211A16"
                      /clone_lib="RPCI-23"
                      /sex="Female"
                      /lab_host="DH10B"
                      /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                      EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                      brain genomic DNA was isolated and  partially digested
                      with a combination of EcoRI and EcoRI Methylase.   Size
                      selected DNA was cloned into the pBACe3.6 vector at the
                      EcoRI  sites.  The ligation products were transformed into
                      DH10B  electrocompetent cells (BRL Life Technologies).  "
BASE COUNT       88 a     49 c     40 g     89 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 17;  Length 266;
  Best Local Similarity   89.5%;  Pred. No. 1.5e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGC 19
              ||||||||| || ||||||
Db         30 TGCAGATTGTGCTATCTGC 12


RESULT 11
AI913095
LOCUS       AI913095               269 bp    mRNA    linear   EST 16-DEC-1999
DEFINITION  tz88a12.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2295646 3',
            mRNA sequence.
ACCESSION   AI913095
VERSION     AI913095.1  GI:5632950
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 269)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
```

```
            Insert Length: 527   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 246.
FEATURES                Location/Qualifiers
     source             1. .269
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:2295646"
                        /clone_lib="NCI_CGAP_Kid11"
                        /lab_host="DH10B"
                        /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                        a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                        Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                        prepared, and ss circles were made in vitro. Following HAP
                        purification, this DNA was used as tracer in a subtractive
                        hybridization reaction. The driver was PCR-amplified cDNAs
                        from a pool of 5,000 clones made from the same library
                        (cloneIDs 1322376-1323911, 1456007-1456775, and
                        1500552-1502855). Subtraction by Bento Soares and M.
                        Fatima Bonaldo. "
BASE COUNT       29 a    104 c     91 g     45 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 9;  Length 269;
  Best Local Similarity   89.5%;  Pred. No. 1.5e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCAGATTGCGCAATCTGC 19
              |||||| |||||| |||||
Db         62 TGCAGACTGCGCAGTCTGC 80


RESULT 12
AI913095/c
LOCUS       AI913095                269 bp    mRNA    linear   EST 16-DEC-1999
DEFINITION  tz88a12.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2295646 3',
            mRNA sequence.
ACCESSION   AI913095
VERSION     AI913095.1  GI:5632950
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 269)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 527   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 246.
FEATURES                Location/Qualifiers
     source             1. .269
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:2295646"
                        /clone_lib="NCI_CGAP_Kid11"
                        /lab_host="DH10B"
                        /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                        a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                        Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                        prepared, and ss circles were made in vitro. Following HAP
                        purification, this DNA was used as tracer in a subtractive
                        hybridization reaction. The driver was PCR-amplified cDNAs
                        from a pool of 5,000 clones made from the same library
                        (cloneIDs 1322376-1323911, 1456007-1456775, and
                        1500552-1502855). Subtraction by Bento Soares and M.
                        Fatima Bonaldo. "
BASE COUNT       29 a    104 c     91 g     45 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 9;  Length 269;
  Best Local Similarity   89.5%;  Pred. No. 1.5e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCAGATTGCGCAATCTGCA 20
              ||||| |||||| ||||||
Db         80 GCAGACTGCGCAGTCTGCA 62


RESULT 13
BB071586
LOCUS       BB071586                301 bp    mRNA    linear   EST 27-JUN-2000
DEFINITION  BB071586 RIKEN full-length enriched, 15 days embryo male testis Mus
            musculus cDNA clone B030498B09 3', mRNA sequence.
ACCESSION   BB071586
VERSION     BB071586.1  GI:8581584
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 301)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES                Location/Qualifiers
     source             1. .301
                        /organism="Mus musculus"
                        /strain="C57BL/6J"
```

/db_xref="taxon:10090"
/clone="8030498B09"
/clone_lib="RIKEN full-length enriched, 15 days embryo male testis"
/sex="male"
/tissue_type="testis"
/dev_stage="15 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 185.0 Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
BASE COUNT 77 a 78 c 54 g 92 t
ORIGIN

Query Match 79.0%; Score 15.8; DB 10; Length 301;
Best Local Similarity 89.5%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGCAGATTGCGCAATCTGC 19
||||||||| | |||||||
Db 224 TGCAGATTGAGAAATCTGC 242

RESULT 14
BB071586/c
LOCUS BB071586 301 bp mRNA linear EST 27-JUN-2000
DEFINITION BB071586 RIKEN full-length enriched, 15 days embryo male testis Mus musculus cDNA clone 8030498B09 3', mRNA sequence.
ACCESSION BB071586
VERSION BB071586.1 GI:8581584
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 301)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.
FEATURES Location/Qualifiers
source 1. .301
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="8030498B09"
/clone_lib="RIKEN full-length enriched, 15 days embryo male testis"
/sex="male"
/tissue_type="testis"
/dev_stage="15 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 185.0 Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
BASE COUNT 77 a 78 c 54 g 92 t
ORIGIN

Query Match 79.0%; Score 15.8; DB 10; Length 301;
Best Local Similarity 89.5%; Pred. No. 1.6e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GCAGATTGCGCAATCTGCA 20
||||||| | |||||||||
Db 242 GCAGATTTCTCAATCTGCA 224

RESULT 15
BU026495
LOCUS BU026495 393 bp mRNA linear EST 23-AUG-2002
DEFINITION QHG17A10.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA clone QHG17A10, mRNA sequence.
ACCESSION BU026495
VERSION BU026495.1 GI:22462015
KEYWORDS EST.
SOURCE common sunflower.
ORGANISM Helianthus annuus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.
REFERENCE 1 (bases 1 to 393)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison

```
           ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
           Church,S., Jackson,L. and Bradford,K.
  TITLE    Lettuce and Sunflower ESTs from the Compositae Genome Project
           http://compgenomics.ucdavis.edu/
  JOURNAL  Unpublished (2002)
COMMENT    Contact: Alexander Kozik [R.W.Michelmore]
           Department of Vegetable Crops, R.W.Michelmore Lab
           University of California at Davis (UCD)
           Asmundson Hall, UCD, Davis, CA 95616, USA
           Tel: 1-(530)-742-1742
           Fax: 1-(530)-752-9659
           Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
           singleton, see http://cgpdb.ucdavis.edu/ for details.
           Plate: QHG17  row: A  column: 10.
FEATURES             Location/Qualifiers
     source          1. .393
                     /organism="Helianthus annuus"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHG17A10"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /lab_host="E.coli"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QH_EFGHJ sunflower RHA280
                     TAG_TISSUE=hulls
                     TAG_SEQ=GCTAGTCGGG"
BASE COUNT      101 a    108 c     85 g     99 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 393;
  Best Local Similarity   89.5%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCAGATTGCGCAATCTGCA 20
              |||||||||| |||| ||
Db        107 GCAGATTGCGCTATCTCCA 125


Search completed: June 26, 2003, 22:12:33
Job time : 1531.13 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 09:43:36 ; Search time 226.087 Seconds
(without alignments)
199.215 Million cell updates/sec

Title: US-09-355-254F-19
Perfect score: 20
Sequence: 1 ctgatttccccgaaatgatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	19	AAV46011	Immune adjuvant ST
2	20	100.0	20	24	AAL39169	Murine Toll-like r
3	20	100.0	524	11	AAQ03455	Recombinant molecu
4	20	100.0	524	11	AAQ03709	Murine interferon
5	20	100.0	524	14	AAQ52649	Promoter and regul
6	18.4	92.0	20	18	AAT88819	Leptin response el
7	18.4	92.0	20	19	AAV38478	IRF-1 gamma interf
8	18.4	92.0	25	21	AAZ92038	STAT5 binding sequ
9	18.4	92.0	669	22	AAH46291	Human interferon r
10	18.4	92.0	669	22	AAH46292	Human interferon r
11	18.4	92.0	669	22	AAH46293	Human interferon r
12	18.4	92.0	700	22	AAH91999	Human inflammatory
13	18	90.0	86	19	AAV59503	Upstream primer fo
14	18	90.0	86	19	AAV34146	Upstream primer fo
15	18	90.0	86	19	AAV34278	Upstream primer fo
16	18	90.0	86	19	AAV69603	Upstream primer fo
17	18	90.0	86	20	AAZ32073	Gamma activation s
18	18	90.0	86	20	AAZ19852	SV40 early promote
19	18	90.0	86	20	AAZ24803	Upstream primer fo
20	18	90.0	86	20	AAZ09776	Synthetic GAS-cont
21	18	90.0	86	20	AAZ10678	PCR primer used to
22	18	90.0	86	20	AAZ00402	Human GAS promoter
23	18	90.0	86	20	AAZ00794	SV40 early promote
24	18	90.0	86	20	AAZ06211	Upstream primer fo
25	18	90.0	86	20	AAX97908	Upstream primer fo
26	18	90.0	86	20	AAX79003	Upstream primer fo
27	18	90.0	86	20	AAX28669	Nucleotide sequenc
28	18	90.0	86	20	AAX84925	Upstream primer fo
29	18	90.0	86	20	AAX35893	PCR primer used to
30	18	90.0	86	20	AAX37361	Human GAS-containi
31	18	90.0	86	20	AAX37443	Synthetic GAS-cont
32	18	90.0	86	20	AAX27303	Upstream primer fo
33	18	90.0	86	20	AAX51693	5' PCR primer used
34	18	90.0	86	20	AAX30175	Upstream primer fo
35	18	90.0	86	20	AAX22203	Upstream primer fo
36	18	90.0	86	20	AAX22103	Upstream primer fo
37	18	90.0	86	20	AAX30309	5' PCR primer used
38	18	90.0	86	20	AAX20404	Upstream primer fo
39	18	90.0	86	20	AAX16170	SV40 early promote
40	18	90.0	86	20	AAX04303	Upstream primer fo
41	18	90.0	86	20	AAX00603	Upstream primer fo
42	18	90.0	86	20	AAV08847	Primer for DNA enc
43	18	90.0	86	20	AAV84403	Upstream primer fo
44	18	90.0	86	21	AAD02079	5' PCR primer to g
45	18	90.0	86	21	AAD02231	SV40 promoter sequ

ALIGNMENTS

RESULT 1
AAV46011
ID AAV46011 standard; DNA; 20 BP.
XX
AC AAV46011;
XX
DT 16-OCT-1998 (first entry)
XX
DE Immune adjuvant STAT4.
XX
KW Immune system; adjuvant; vaccine; cancer; prophylactic; pathogenicity;
KW modulator; tolerance; regulator; helper cell; antigen; immunoglobulin;
KW Ig class; autoimmune response; T-cell; B-cell; tumour; ss.
XX
OS Class Bacteria.
XX
PN EP855184-A1.
XX
PD 29-JUL-1998.
XX
PF 23-JAN-1997; 97EP-0101019.
XX
PR 23-JAN-1997; 97EP-0101019.
XX
PA (HEEG/) HEEG K.
PA (LIPF/) LIPFORD G B.
PA (WAGN/) WAGNER H.
XX
PI Heeg K, Lipford GB, Wagner H;
XX
DR WPI; 1998-389630/34.
XX

PT Antigenic composition comprises polynucleotide fragment and antigen
PT - used as vaccine to treat or prevent e.g. cancer or pathogen
PT infections and to modulate immune response e.g. tolerance break and
PT regulation of TH1/TH2 cells
XX
PS Example 5; Page 9; 28pp; English.
XX
CC AAV45993-V46019 are fragments of bacterial polynucleotides which are
CC used as immune adjuvants for inclusion into vaccines to treat cancer and
CC for prophylaxis and/or treatment of conditions caused by pathogenic
CC micro-organisms. The polynucleotide is used for modulation of an immune
CC response and the modulation is selected from the group break of
CC tolerance, regulation of TH1/TH2 helper cell responses, switch of Ig
CC classes, treatment of autoimmune responses and induction of tolerances.
CC DNA oligomers are used to enhance the reactivity of immune cells to
CC viral, bacterial and parasitic antigens, to break tolerance in anergic T
CC and B cells e.g. against tumour antigens, as adjuvants in vaccination
CC against tumour-defined antigens and immunostimulatory substances in an
CC immune response against tumours and to suppress immune reactions of the
CC innate and acquired immune system. The composition is inexpensive and
CC stable and does not cause lethal shock, which happens with prior art.
CC bacterial sequences.
XX
SQ Sequence 20 BP; 5 A; 5 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGATTTCCCCGAAATGATG 20
||||||||||||||||||||
Db 1 CTGATTTCCCCGAAATGATG 20

RESULT 2
AAL39169
ID AAL39169 standard; DNA; 20 BP.
XX
AC AAL39169;
XX
DT 05-SEP-2002 (first entry)
XX
DE Murine Toll-like receptor related CpG DNA SEQ ID No 44.
XX
KW Murine Toll-like receptor; TLR9; TLR7; TLR8; ISNA; ds.
XX
OS Unidentified.
XX
PN WO200222809-A2.
XX
PD 21-MAR-2002.
XX
PF 17-SEP-2001; 2001WO-US29229.
XX
PR 15-SEP-2000; 2000US-233035P.
PR 23-JAN-2001; 2001US-263657P.
PR 17-MAY-2001; 2001US-291726P.
PR 22-JUN-2001; 2001US-300210P.
XX
PA (COLE-) COLEY PHARM GMBH.
XX
PI Bauer S, Lipford G, Wagner H;
XX
DR WPI; 2002-393964/42.
XX
PT New isolated murine Toll-like receptor (TLR)9, TLR7, TLR8 polypeptides,
PT useful for identifying species specificity of immunostimulatory nucleic
PT acid and identifying immunostimulatory nucleic acids -
XX
PS Disclosure; Page 76; 195pp; English.
XX
CC The invention relates to isolated murine Toll-like receptors (TLR)9,
CC TLR7 and TLR8 polypeptides. These polypeptides comprise fully defined
CC sequences of 1032, 1050 or 1032 amino acids as given in specification, or
CC their fragments, where TLR9, TLR7 and TLR8 polypeptides or their
CC fragments have an amino acid sequence which is identical to human TLR9,
CC TLR7 or TLR8 polypeptides or their fragment except for at least one amino
CC acid of a murine TLR polypeptide. The isolated nucleic acids of the
CC invention are useful for inhibiting TLR9 signalling activity in a cell.
CC TLR7, TLR8 and TLR9 polypeptides are useful for identifying nucleic acid
CC molecules which interact with a TLR polypeptide or its fragment. The
CC TLR7, TLR8 or TLR9 polypeptides are also useful for identifying ISNA. The
CC TLR7, TLR8 and TLR9 polypeptides are also useful for comparing TLR9
CC signalling activity of a test compound (that is not a nucleic acid, and
CC is a polypeptide or a part of a combinatorial library of compounds) with
CC an ISNA. The TLR7, TLR8 and TLR9 polypeptides are also useful for
CC identifying species specificity of an ISNA. The isolated nucleic acids of
CC the invention are useful as probes or primers. This polynucleotide
CC sequence represents DNA relating to the isolated Toll-like receptors of
CC the invention.
XX
SQ Sequence 20 BP; 5 A; 5 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGATTTCCCCGAAATGATG 20
||||||||||||||||||||
Db 1 CTGATTTCCCCGAAATGATG 20

RESULT 3
AAQ03455
ID AAQ03455 standard; DNA; 524 BP.
XX
AC AAQ03455;
XX
DT 23-JUL-1990 (first entry)
XX
DE Recombinant molecule encoding protein having IRF-1 activity.
XX
KW Regulatory factor-1, interferon-beta gene, cis-elements; ss.
XX
OS Eukaryotic.
XX
FH Key Location/Qualifiers
FT GC_signal 70..76
FT /*tag= a
FT /label=GC Box 1
FT GC_signal 92..98
FT /*tag= b
FT /label=GC Box 2
FT CAAT_signal 201..207
FT /*tag= c
FT misc_signal 278..280
FT /*tag= d
FT /label=Minor Cap site
FT misc_signal 299..301
FT /*tag= e
FT /label=Major Cap site
FT CDS 330..524
FT /*tag= f
FT /label=pIRF-1
XX
PN EP355202-A.
XX
PD 28-FEB-1990.
XX
PF 24-NOV-1988; 88EP-0119602.
XX
PR 24-NOV-1988; 88EP-0113793.
XX
PA (TANI/) TANIGUCHI T.
XX
PI Taniguchi T;

```
XX
DR   WPI; 1990-060144/09.
XX
PT   Interferon regulatory factor-1 - which is active in virus induced
PT   activation of interferon-beta gene transcription by interacting with
PT   cis-elements.
XX
PS   Claim 13; page 46; 65pp; English.
XX
CC   Recombinant molecule containing promoter and regulatory sequence; can
CC   be used for the production of IRF-1. IRF-1 plays an essential role
CC   in virus-induced activation of interferon-beta gene transcription by
CC   interacting with the cis-elements. The recombinant molecule can
CC   also be designed for expression of pharmaceutically active proteins
CC   such as e.g. cytokine or plasminogen activator.
CC   See also AAQ03452,-51 and -53; and EP-355190-A.
XX
SQ   Sequence 524 BP; 90 A; 166 C; 185 G; 83 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 11;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              ||||||||||||||||||||
Db        171 CTGATTTCCCCGAAATGATG 190


RESULT 4
AAQ03709
ID   AAQ03709 standard; DNA; 524 BP.
XX
AC   AAQ03709;
XX
DT   03-AUG-1990  (first entry)
XX
DE   Murine interferon regulatory factor-1 promoter and regulatory sequence.
XX
KW   Interferon regulatory factor; IRF-1 beta-gene; interferon;
KW   murine; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CAAT_signal     201..207
FT                   /*tag=  a
FT   GC_signal       70..76
FT                   /*tag=  b
FT   GC_signal       92..98
FT                   /*tag=  c
FT   misc_feature    77..79
FT                   /*tag=  d
FT                   /label=Minor cap site.
FT   misc_feature    298..300
FT                   /*tag=  e
FT                   /label=Major cap site.
FT   CDS             300..523
FT                   /*tag=  f
XX
PN   EP359998-A.
XX
PD   28-MAR-1990.
XX
PF   17-AUG-1989;   89EP-0115158.
XX
PR   24-AUG-1988;   88EP-0113793.
PR   24-NOV-1988;   88EP-0119602.
XX
PA   (TANI/) TANIGUCHI T.
XX
PI   Taniguchi T,  Fuchita T;
XX
DR   WPI; 1990-092658/13.
XX
PT   Recombinant interferon regulator factor-1 -
PT   which plays an essential role in virus-induced activation of
PT   interferon-beta gene transcription.
XX
PS   Claim 10; Fig 7; 55pp; English.
XX
CC   Promoter and regulator sequence for IFN regulator factor-1.
XX
SQ   Sequence 524 BP; 90 A; 166 C; 185 G; 83 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 11;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              ||||||||||||||||||||
Db        171 CTGATTTCCCCGAAATGATG 190


RESULT 5
AAQ52649
ID   AAQ52649 standard; DNA; 524 BP.
XX
AC   AAQ52649;
XX
DT   26-MAY-1994  (first entry)
XX
DE   Promoter and regulatory sequence used in recombinant construct.
XX
KW   Interferon; interferon-beta; regulation; gene expression;
KW   regulatory element; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   GC_signal       70..76
FT                   /*tag=  a
FT                   /note= "GC Box 1."
FT   GC_signal       92..98
FT                   /*tag=  b
FT                   /note= "GC box 2."
FT   CAAT_signal     201..207
FT                   /*tag=  c
FT   misc_feature    278..280
FT                   /*tag=  d
FT                   /note= "Minor cap site."
FT   misc_feature    299..301
FT                   /*tag=  e
FT                   /note= "Major cap site."
XX
PN   EP571743-A.
XX
PD   01-DEC-1993.
XX
PF   17-AUG-1989;   89EP-0115158.
XX
PR   24-AUG-1988;   88EP-0113793.
PR   24-NOV-1988;   88EP-0119602.
XX
PA   (TANI/) TANIGUCHI T.
XX
PI   Fujita T,  Taniguchi T;
XX
DR   WPI; 1993-378709/48.
XX
PT   Interferon regulatory factor - useful for controlling expression
PT   of interferon genes
XX
PS   Disclosure; Page 8; 45pp; English.
XX
CC   A protein which binds to the repeated oligomer sequence AAGTGA
CC   and regulatory upstream elements of the human interferon regulatory
```

```
CC   factor (IRF)-beta gene (See AAR44217) can be used to regulate the
CC   expression of interferon-beta.  In a recombinant molecule, the
CC   sequence encoding this protein can be placed under the control of
CC   this promoter and regulatory sequence.
XX
SQ   Sequence 524 BP; 90 A; 166 C; 185 G; 83 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 14;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              ||||||||||||||||||||
Db        171 CTGATTTCCCCGAAATGATG 190


RESULT 6
AAT88819
ID   AAT88819 standard; DNA; 20 BP.
XX
AC   AAT88819;
XX
DT   12-MAY-1998  (first entry)
XX
DE   Leptin response element IRF-1 derived gamma interferon activation seq.
XX
KW   Leptin response element; IRF-1; gamma interferon activation sequence;
KW   detection; obesity; diabetes; infertility; cachexia; anorexia;
KW   human; Ob-r; ss.
XX
OS   Synthetic.
XX
PN   WO9740380-A1.
XX
PD   30-OCT-1997.
XX
PF   18-APR-1997;   97WO-US06505.
XX
PR   15-NOV-1996;   96US-0031002.
PR   22-APR-1996;   96US-0016051.
PR   06-JUN-1996;   96GB-0011785.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Chen F,  Cully DF,  Hess JW,  Qureshi SA,  Rosenblum CI;
PI   Tota MR,  Van Der Ploeg L;
XX
DR   WPI; 1997-536003/49.
XX
PT   Detecting and quantifying leptin and related compounds - used e.g.
PT   to assay recombinant leptin products, to identify agonists, etc
XX
PS   Claim 3; Page 12; 28pp; English.
XX
CC   The present sequence represents a leptin response element IRF-1 derived
CC   gamma-interferon activation sequence. A method has been developed for
CC   determining the presence of a leptin-receptor binding component (A) in
CC   a sample. The method involves: (i) treating the sample with cells
CC   containing: (a) nucleic acid (I) comprising a reporter gene (RG) linked
CC   to a promoter that includes at least one leptin response element (LRE);
CC   and (b) nucleic acid (II) encoding a leptin receptor (LR); and (ii)
CC   determining if transcription of RG occurs. The method is used to assay
CC   activity in recombinant leptin preparations and to identify leptin
CC   (ant)agonists which are potentially useful for treating obesity,
CC   diabetes and infertility (agonists) or cachexia or anorexia
CC   (antagonists). Cells transformed with the vector are used to confirm
CC   activity of putative LRE and to determine their leptin-binding activity
CC   (by their response to added leptin). Also the binding of different LR
CC   can be compared. The method is an alternative to current animal tests
CC   on ob/ob mice.
XX
SQ   Sequence 20 BP; 5 A; 6 C; 4 G; 5 T; 0 other;

  Query Match             92.0%;  Score 18.4;  DB 18;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 7.4;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              |||||||||||||||||| |
Db          1 CTGATTTCCCCGAAATGACG 20


RESULT 7
AAV38478
ID   AAV38478 standard; DNA; 20 BP.
XX
AC   AAV38478;
XX
DT   12-OCT-1998  (first entry)
XX
DE   IRF-1 gamma interferon activation sequence.
XX
KW   ob-receptor; hypothalamus; obesity; leptin; diabetes; infertility;
KW   anorexia; cachexia; ss.
XX
OS   Synthetic.
XX
PN   WO9824881-A1.
XX
PD   11-JUN-1998.
XX
PF   26-NOV-1997;   97WO-US22165.
XX
PR   02-DEC-1996;   96US-0032367.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Fong TM,  Huang RC,  Van Der Ploeg L;
XX
DR   WPI; 1998-333304/29.
XX
PT   New mutant ob receptor(s) - used to develop products for drug
PT   screening and for gene therapy for weight control, e.g. obesity or
PT   anorexia
XX
PS   Disclosure; Page 9; 27pp; English.
XX
CC   The IRF-1 derived gamma interferon activation sequence is used in
CC   in the construction of a reporter gene for use in determining whether an
CC   OB-R ligand is present in a sample.  The ob-receptor (OB-R), a member of
CC   the cytokine receptor family is transcribed in the hypothalamus and is
CC   involved in obesity.  Mutants lacking a functional first or second CK-F3
CC   module or a functional intracellular domain can be used in assays for the
CC   detection of ligands, agonists, antagonists and ligand mimetics.  The
CC   leptin agonists identified can be used in situations where leptin
CC   insufficiency causes obesity, diabetes or infertility.  The leptin
CC   antagonists identified can be used in the treatment of anorexia and
CC   cachexia.  The mutant receptor nucleic acids can also be used in gene
CC   therapy for weight control, e.g. for treating obesity or anorexia.
XX
SQ   Sequence 20 BP; 5 A; 6 C; 4 G; 5 T; 0 other;

  Query Match             92.0%;  Score 18.4;  DB 19;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 7.4;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              |||||||||||||||||| |
Db          1 CTGATTTCCCCGAAATGACG 20


RESULT 8
AAZ92038
ID   AAZ92038 standard; DNA; 25 BP.
XX
AC   AAZ92038;
```

```
XX
DT   08-JUN-2000  (first entry)
XX
DE   STAT5 binding sequence oligonucleotide IRF-1 GAS.
XX
KW   STAT5 protein; signal transducer and activator of transcription 5;
KW   protein binding sequence; transcription factor modulator; inhibitor;
KW   malignant cell removal; proliferative malignancy; neoplastic disease;
KW   immunological disorder; inflammatory disorder; therapy; ds.
XX
OS   Synthetic.
XX
PN   WO200006696-A2.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-US17366.
XX
PR   30-JUL-1998;   98US-0094695.
XX
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Zuckerman KS,  Liu RY;
XX
DR   WPI; 2000-195281/17.
XX
PT   Therapeutic agent for treating transcription factor-related illnesses
PT   such as proliferative malignancies, comprises an oligonucleotide for
PT   regulating transcription factor function -
XX
PS   Claim 15; Page 34; 43pp; English.
XX
CC   This sequence represents a STAT5 (signal transducer and activator of
CC   transcription 5) protein binding sequence. The invention relates to a
CC   therapeutic agent comprising an effective amount of an oligonucleotide
CC   (I) for modulating the function of transcription factors and a
CC   pharmaceutical acceptable carrier. The oligonucleotides can be used in a
CC   method of removing malignant cells in vitro. The oligonucleotides can be
CC   used in compositions to inhibit transcription factors in illnesses where
CC   transcription factors play a role, especially proliferative malignancies,
CC   neoplastic diseases, and immunological and inflammatory disorders.
XX
SQ   Sequence 25 BP; 6 A; 8 C; 6 G; 5 T; 0 other;

  Query Match             92.0%;  Score 18.4;  DB 21;  Length 25;
  Best Local Similarity   95.0%;  Pred. No. 7.6;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              |||||||||||||||||| |
Db          3 CTGATTTCCCCGAAATGACG 22


RESULT 9
AAH46291
ID   AAH46291 standard; DNA; 669 BP.
XX
AC   AAH46291;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human interferon regulatory factor-1 promoter region, wild-type allele.
XX
KW   Human; interferon regulatory factor-1; IRF-1; promoter;
KW   upstream region; genotyping; polymorphism; hepatitis C virus;
KW   HCV infection; interferon therapy efficacy; IFN; RFLP analysis;
KW   restriction fragment length polymorphism; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   allele          replace (108, C)
FT                   /*tag=  a
FT                   /note= "This nucleotide substitution is present in
FT                          the IRF-1 alleles found in HCC-T/HCC-M (IFN-
FT                          sensitive) and PLC/PRF/5 (IFN insensitive) liver
FT                          cancer cells"
FT   allele          replace (196, A)
FT                   /*tag=  b
FT                   /note= "This nucleotide substitution is present
FT                          only in the IRF-1 allele found in PLC/PRF/5
FT                          (IFN insensitive) liver cancer cells"
XX
PN   JP2001136973-A.
XX
PD   22-MAY-2001.
XX
PF   16-NOV-1999;   99JP-0324975.
XX
PR   16-NOV-1999;   99JP-0324975.
XX
PA   (SAKA ) OTSUKA PHARM CO LTD.
XX
DR   WPI; 2001-460211/50.
XX
PT   Detection of abnormal human interferon regulatory factor-1 (IRF-1) gene
PT   -
XX
PS   Example 2; Fig 1-2; 8pp; Japanese.
XX
CC   The invention relates to a method for the detection of an abnormal
CC   allele of the human interferon regulatory factor-1 (IRF-1) gene. The
CC   abnormal allele (AAH46293) is present in PLC/PRF/5 liver cancer cells
CC   and contains a G to A substitution at position 196 of the IRF-1 promoter
CC   region (normal alleles given in AAH46293 and AAH46294). The abnormal
CC   allele confers an insensitivity to the effects of interferon (IFN).
CC   In the method of the invention, the presence or absence of adenine
CC   at position 196 of the IRF-1 promoter is detected using procedures such
CC   as restriction fragment length polymorphism (RFLP) analysis. Prior to
CC   analysis, an IRF-1 gene fragment containing the polymorphic site can
CC   optionally be prepared (e.g., by PCR). The invention also discloses the
CC   use of IRF-1 gene fragments as probes to detect the A polymorphism. The
CC   method of the invention is used to genotype a patient with hepatitis C
CC   virus (HCV) infection in order to predict whether interferon therapy
CC   will be effective. The present sequence represents the wild-type
CC   allele of the human IRF-1 promoter region (also given in GenBank
CC   accession number X53095) which was used in RFLP analysis.
XX
SQ   Sequence 669 BP; 97 A; 234 C; 247 G; 91 T; 0 other;

  Query Match             92.0%;  Score 18.4;  DB 22;  Length 669;
  Best Local Similarity   95.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              |||||||||||||||||| |
Db        367 CTGATTTCCCCGAAATGACG 386


RESULT 10
AAH46292
ID   AAH46292 standard; DNA; 669 BP.
XX
AC   AAH46292;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human interferon regulatory factor-1 promoter region, HCC-T/HCC-M allele.
XX
KW   Human; interferon regulatory factor-1; IRF-1; promoter;
KW   upstream region; genotyping; polymorphism; hepatitis C virus;
KW   HCV infection; interferon therapy efficacy; IFN; RFLP analysis;
KW   restriction fragment length polymorphism; HCC-T allele; HCC-M; ds.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   allele          replace (108, T)
FT                   /*tag= a
FT   allele          replace (196, A)
FT                   /*tag= b
FT                   /note= "This nucleotide substitution is additionally
FT                          present in the IRF-1 allele found in PLC/PRF/5
FT                          (IFN insensitive) liver cancer cells"
XX
PN   JP2001136973-A.
XX
PD   22-MAY-2001.
XX
PF   16-NOV-1999;   99JP-0324975.
XX
PR   16-NOV-1999;   99JP-0324975.
XX
PA   (SAKA ) OTSUKA PHARM CO LTD.
XX
DR   WPI; 2001-460211/50.
XX
PT   Detection of abnormal human interferon regulatory factor-1 (IRF-1) gene
PT   -
XX
PS   Example 2; Fig 1-2; 8pp; Japanese.
XX
CC   The invention relates to a method for the detection of an abnormal
CC   allele of the human interferon regulatory factor-1 (IRF-1) gene. The
CC   abnormal allele (AAH46293) is present in PLC/PRF/5 liver cancer cells
CC   and contains a G to A substitution at position 196 of the IRF-1 promoter
CC   region (normal alleles given in AAH46293 and AAH46294). The abnormal
CC   allele confers an insensitivity to the effects of interferon (IFN).
CC   In the method of the invention, the presence or absence of adenine
CC   at position 196 of the IRF-1 promoter is detected using procedures such
CC   as restriction fragment length polymorphism (RFLP) analysis. Prior to
CC   analysis, an IRF-1 gene fragment containing the polymorphic site can
CC   optionally be prepared (e.g., by PCR). The invention also discloses the
CC   use of IRF-1 gene fragments as probes to detect the A polymorphism. The
CC   method of the invention is used to genotype a patient with hepatitis C
CC   virus (HCV) infection in order to predict whether interferon therapy
CC   will be effective. The present sequence represents the allele of the
CC   human IRF-1 promoter region found in IFN-sensitive HCC-T and HCC-M
CC   liver cancer cells which was used in RFLP analysis.
XX
SQ   Sequence 669 BP; 97 A; 235 C; 247 G; 90 T; 0 other;

  Query Match             92.0%;  Score 18.4;  DB 22;  Length 669;
  Best Local Similarity   95.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              |||||||||||||||||| |
Db        367 CTGATTTCCCCGAAATGACG 386


RESULT 11
AAH46293
ID   AAH46293 standard; DNA; 669 BP.
XX
AC   AAH46293;
XX
DT   25-SEP-2001  (first entry)
XX
DE   Human interferon regulatory factor-1 promoter region, PLC/PRF/5 allele.
XX
KW   Human; interferon regulatory factor-1; IRF-1; promoter;
KW   upstream region; genotyping; polymorphism; hepatitis C virus;
KW   HCV infection; interferon therapy efficacy; IFN; RFLP analysis;
KW   restriction fragment length polymorphism; PLC/PRF/5 allele; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   allele          replace (108, T)
FT                   /*tag= a
FT   allele          replace (196, G)
FT                   /*tag= b
XX
PN   JP2001136973-A.
XX
PD   22-MAY-2001.
XX
PF   16-NOV-1999;   99JP-0324975.
XX
PR   16-NOV-1999;   99JP-0324975.
XX
PA   (SAKA ) OTSUKA PHARM CO LTD.
XX
DR   WPI; 2001-460211/50.
XX
PT   Detection of abnormal human interferon regulatory factor-1 (IRF-1) gene
PT   -
XX
PS   Example 1; Fig 1-2; 8pp; Japanese.
XX
CC   The invention relates to a method for the detection of an abnormal
CC   allele of the human interferon regulatory factor-1 (IRF-1) gene. The
CC   abnormal allele (AAH46293) is present in PLC/PRF/5 liver cancer cells
CC   and contains a G to A substitution at position 196 of the IRF-1 promoter
CC   region (normal alleles given in AAH46293 and AAH46294). The abnormal
CC   allele confers an insensitivity to the effects of interferon (IFN).
CC   In the method of the invention, the presence or absence of adenine
CC   at position 196 of the IRF-1 promoter is detected using procedures such
CC   as restriction fragment length polymorphism (RFLP) analysis. Prior to
CC   analysis, an IRF-1 gene fragment containing the polymorphic site can
CC   optionally be prepared (e.g., by PCR). The invention also discloses the
CC   use of IRF-1 gene fragments as probes to detect the A polymorphism. The
CC   method of the invention is used to genotype a patient with hepatitis C
CC   virus (HCV) infection in order to predict whether interferon therapy
CC   will be effective. The present sequence represents the allele of the
CC   human IRF-1 promoter region found in IFN-insensitive PLC/PRF/5
CC   liver cancer cells which was used in RFLP analysis.
XX
SQ   Sequence 669 BP; 98 A; 235 C; 246 G; 90 T; 0 other;

  Query Match             92.0%;  Score 18.4;  DB 22;  Length 669;
  Best Local Similarity   95.0%;  Pred. No. 11;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              |||||||||||||||||| |
Db        367 CTGATTTCCCCGAAATGACG 386


RESULT 12
AAH91999
ID   AAH91999 standard; DNA; 700 BP.
XX
AC   AAH91999;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Human inflammatory bowel disease related gene fragment IGR2011a.
XX
KW   Human; inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW   single nucleotide polymorphism; SNP; chromosome 19p13; paternity test;
KW   chromosome 5q31-33; forensic test; gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200142511-A2.
XX
PD   14-JUN-2001.
XX
PF   11-DEC-2000; 2000WO-US33632.
XX
```

PR 10-DEC-1999; 99US-0170257.
PR 10-APR-2000; 2000US-0196046.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PA (ELLI-) ELLIPSIS BIOTHERAPEUTICS CORP.
XX
PI Daly M, Hudson TJ, Lander ES, Rioux J, Siminovitch K;
XX
DR WPI; 2001-367874/38.
XX
PT Testing for the presence of polymorphisms associated with inflammatory
PT bowel disease, using a hybridization assay -
XX
PS Disclosure; Page 89; 463pp; English.
XX
CC The present invention describes a method for detecting the presence of
CC polymorphisms associated with inflammatory bowel diseases such as
CC ulcerative colitis and Crohn's disease. The methods can be used to detect
CC the presence of genetic polymorphisms associated with inflammatory bowel
CC disease and correlating their occurrence with disease states. They may be
CC used in this way for phenotypic correlations, forensics, paternity
CC testing, medicine and genetic analysis. The present sequence is a gene
CC containing a polymorphic site described in the exemplification of the
CC invention.
XX
SQ Sequence 700 BP; 97 A; 255 C; 254 G; 94 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 700;
Best Local Similarity 95.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGATTTCCCCGAAATGATG 20
||||||||||||||||||| |
Db 331 CTGATTTCCCCGAAATGACG 350

RESULT 13
AAV59503
ID AAV59503 standard; DNA; 86 BP.
XX
AC AAV59503;
XX
DT 02-FEB-1999 (first entry)
XX
DE Upstream primer for SV40 promoter sequence.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Synthetic.
OS Simian virus 40.
XX
PN WO9839448-A2.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US04493.
XX
PR 02-OCT-1997; 97US-0061060.
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.

```
PR   22-AUG-1997;   97US-0056884.
PR   22-AUG-1997;   97US-0056886.
PR   22-AUG-1997;   97US-0056887.
PR   22-AUG-1997;   97US-0056888.
PR   22-AUG-1997;   97US-0056889.
PR   22-AUG-1997;   97US-0056892.
PR   22-AUG-1997;   97US-0056893.
PR   22-AUG-1997;   97US-0056894.
PR   22-AUG-1997;   97US-0056903.
PR   22-AUG-1997;   97US-0056908.
PR   22-AUG-1997;   97US-0056909.
PR   22-AUG-1997;   97US-0056910.
PR   22-AUG-1997;   97US-0056911.
PR   05-SEP-1997;   97US-0057650.
PR   05-SEP-1997;   97US-0057669.
PR   05-SEP-1997;   97US-0057761.
PR   12-SEP-1997;   97US-0058785.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI   Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI   Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI   Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
DR   WPI; 1998-506364/43.
XX
PT   New isolated human genes and the secreted polypeptide(s) they encode
PT   - useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders
XX
PS   Example 12; Page 220; 721pp; English.
XX
CC   The invention relates to 186 novel genes and their fragments (nucleic
CC   acid sequences: AAV59511-V59812; amino acid sequences AAW74731-W75026)
CC   which are useful for preventing, treating or ameliorating medical
CC   conditions e.g. by protein or gene therapy. Also, pathological
CC   conditions can be diagnosed by determining the amount of the new
CC   polypeptides in a sample or by determining the presence of mutations in
CC   the new polynucleotides. Specific uses are described for each of the 186
CC   polynucleotides, based on which tissues they are most highly expressed in
CC   (see AAV59511 for described uses). The genes can be used to generate
CC   fusion proteins by linking to the gene to a sequence encoding human
CC   immunoglobulin Fc portion (AAV59502) for increasing the stability of the
CC   fused protein as compared to the secreted protein only. Genes encoding
CC   the secreted proteins can be used for high-throughput assays for
CC   biological activities. Expression of the genes can be driven by a range
CC   of promoter active in eukaryotic cells. Primers AAV59503-V59504 are used
CC   to amplify the simian virus 40 (SV40) promoter (AAV59505) to generate a
CC   construct for identifying proteins involved in signal transduction which
CC   bind the gamma activation site (GAS) in a similar manner to the Jaks-STAT
CC   pathways.
XX
SQ   Sequence 86 BP; 23 A; 26 C; 13 G; 24 T; 0 other;

  Query Match             90.0%;  Score 18;  DB 19;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57


RESULT 14
AAV34146
ID   AAV34146 standard; DNA; 86 BP.
XX
AC   AAV34146;
XX
DT   02-FEB-1999  (first entry)
XX
DE   Upstream primer for SV40 promoter sequence.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Synthetic.
OS   Simian virus 40.
XX
PN   WO9839446-A2.
XX
PD   11-SEP-1998.
XX
PF   06-MAR-1998;   98WO-US04492.
XX
PR   07-MAR-1997;   97US-0038621.
PR   07-MAR-1997;   97US-0040161.
PR   07-MAR-1997;   97US-0040162.
PR   07-MAR-1997;   97US-0040163.
PR   07-MAR-1997;   97US-0040333.
PR   07-MAR-1997;   97US-0040334.
PR   07-MAR-1997;   97US-0040336.
PR   07-MAR-1997;   97US-0040626.
PR   11-APR-1997;   97US-0043311.
PR   11-APR-1997;   97US-0043312.
PR   11-APR-1997;   97US-0043313.
PR   11-APR-1997;   97US-0043314.
PR   11-APR-1997;   97US-0043315.
PR   11-APR-1997;   97US-0043568.
PR   11-APR-1997;   97US-0043569.
PR   11-APR-1997;   97US-0043576.
PR   11-APR-1997;   97US-0043578.
PR   11-APR-1997;   97US-0043580.
PR   11-APR-1997;   97US-0043669.
PR   11-APR-1997;   97US-0043670.
PR   11-APR-1997;   97US-0043671.
PR   11-APR-1997;   97US-0043672.
PR   11-APR-1997;   97US-0043674.
PR   23-MAY-1997;   97US-0047492.
PR   23-MAY-1997;   97US-0047500.
PR   23-MAY-1997;   97US-0047501.
PR   23-MAY-1997;   97US-0047502.
PR   23-MAY-1997;   97US-0047503.
PR   23-MAY-1997;   97US-0047581.
PR   23-MAY-1997;   97US-0047582.
PR   23-MAY-1997;   97US-0047583.
PR   23-MAY-1997;   97US-0047584.
PR   23-MAY-1997;   97US-0047585.
PR   23-MAY-1997;   97US-0047586.
PR   23-MAY-1997;   97US-0047587.
PR   23-MAY-1997;   97US-0047588.
PR   23-MAY-1997;   97US-0047589.
PR   23-MAY-1997;   97US-0047590.
PR   23-MAY-1997;   97US-0047592.
PR   23-MAY-1997;   97US-0047593.
PR   23-MAY-1997;   97US-0047594.
PR   23-MAY-1997;   97US-0047595.
PR   23-MAY-1997;   97US-0047596.
PR   23-MAY-1997;   97US-0047597.
PR   23-MAY-1997;   97US-0047598.
PR   23-MAY-1997;   97US-0047599.
PR   23-MAY-1997;   97US-0047600.
PR   23-MAY-1997;   97US-0047601.
PR   23-MAY-1997;   97US-0047612.
PR   23-MAY-1997;   97US-0047613.
PR   23-MAY-1997;   97US-0047614.
PR   23-MAY-1997;   97US-0047615.
PR   23-MAY-1997;   97US-0047617.
PR   23-MAY-1997;   97US-0047618.
```

PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057761.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Graves KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
DR WPI; 1998-609887/51.
XX
PT New isolated human genes and the secreted polypeptides they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Example 12; Page 139; 447pp; English.
XX
CC The invention relates to 70 novel genes and their fragments (nucleic
CC acid sequences: AAV34154-V34276; amino acid sequences AAW75057-W75179)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 70
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV34154 for described uses). The genes can be used to generate
CC fusion proteins by linking to the gene to a sequence encoding human
CC immunoglobulin Fc portion (AAV34145) for increasing the stability of the
CC fused protein as compared to the secreted protein only. Genes encoding
CC the secreted proteins can be used for high-throughput assays for
CC biological activities. Expression of the genes can be driven by a range
CC of promoter active in eukaryotic cells. Primers AAV34146-V34147 are used
CC to amplify the simian virus 40 (SV40) promoter (AAV34148) to generate a
CC construct for identifying proteins involved in signal transduction which
CC bind the gamma activation site (GAS) in a similar manner to the Jaks-STAT
CC pathways.
XX
SQ Sequence 86 BP; 23 A; 26 C; 13 G; 24 T; 0 other;

Query Match 90.0%; Score 18; DB 19; Length 86;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TGATTTCCCCGAAATGAT 19
||||||||||||||||||
Db 40 TGATTTCCCCGAAATGAT 57

RESULT 15
AAV34278
ID AAV34278 standard; DNA; 86 BP.
XX
AC AAV34278;
XX
DT 29-JAN-1999 (first entry)
XX
DE Upstream primer for SV40 promoter sequence.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Synthetic.
OS Simian virus 40.
XX
PN WO9840483-A2.
XX
PD 17-SEP-1998.
XX
PF 12-MAR-1998; 98WO-US04858.
XX
PR 19-DEC-1997; 97US-0068368.
PR 14-MAR-1997; 97US-0040710.
PR 14-MAR-1997; 97US-0040762.
PR 30-MAY-1997; 97US-0048100.
PR 30-MAY-1997; 97US-0048189.
PR 30-MAY-1997; 97US-0048357.
PR 30-MAY-1997; 97US-0050934.
PR 06-JUN-1997; 97US-0048970.
PR 05-SEP-1997; 97US-0057765.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ferrie AM, Fischer CL, Gentz RL, Greene JM, Kyaw H;
PI Li H, Li Y, Moore PA, Rosen CA, Ruben SM, Soppet DR;
PI Wei YF, Young PE, Zeng Z;
XX
DR WPI; 1998-520811/44.
XX
PT Isolated human poly:nucleotide(s) encoding secretory peptide(s) -
PT used to develop products for the diagnosis and treatment of e.g.
PT inflammation, cancers, CNS disorders or immune system disorders
XX
PS Example 12; Page 89; 201pp; English.
XX
CC The invention relates to 28 novel genes and their fragments (nucleic
CC acid sequences: AAV34286-V34325; amino acid sequences AAW75196-W75235)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 28
CC polynucleotides, based on which tissues they are most highly expressed in

CC (see AAV34286 for described uses). The genes can be used to generate
CC fusion proteins by linking to the gene to a sequence encoding human
CC immunoglobulin Fc portion (AAV34277) for increasing the stability of the
CC fused protein as compared to the secreted protein only. Genes encoding
CC the secreted proteins can be used for high-throughput assays for
CC biological activities. Expression of the genes can be driven by a range
CC of promoter active in eukaryotic cells. Primers AAV34278-V34279 are used
CC to amplify the simian virus 40 (SV40) promoter (AAV34280) to generate a
CC construct for identifying proteins involved in signal transduction which
CC bind the gamma activation site (GAS) in a similar manner to the Jaks-STAT
CC pathways.
XX
SQ Sequence 86 BP; 23 A; 26 C; 13 G; 24 T; 0 other;

Query Match 90.0%; Score 18; DB 19; Length 86;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TGATTTCCCCGAAATGAT 19
||||||||||||||||||
Db 40 TGATTTCCCCGAAATGAT 57

Search completed: June 26, 2003, 12:16:31
Job time : 228.158 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 10:58:25 ; Search time 30.2174 Seconds
(without alignments)
202.980 Million cell updates/sec

Title: US-09-355-254F-19
Perfect score: 20
Sequence: 1 ctgatttccccgaaatgatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	20	100.0	21	1	US-08-394-191-12	Sequence 12, Appl
c	2	20	100.0	21	1	US-08-458-364-12	Sequence 12, Appl
	3	18.4	92.0	20	3	US-08-837-635-1	Sequence 1, Appli
	4	18	90.0	86	3	US-09-106-182-16	Sequence 16, Appl
	5	18	90.0	86	4	US-09-227-357-3	Sequence 3, Appli
	6	18	90.0	86	4	US-09-280-839-6	Sequence 6, Appli
	7	18	90.0	86	4	US-09-411-977-18	Sequence 18, Appl
	8	18	90.0	86	4	US-09-479-729B-23	Sequence 23, Appl
	9	18	90.0	86	4	US-09-257-179-3	Sequence 3, Appli
	10	18	90.0	86	4	US-09-149-476-3	Sequence 3, Appli
	11	18	90.0	86	4	US-09-288-143-3	Sequence 3, Appli
	12	18	90.0	86	4	US-09-487-792-25	Sequence 25, Appl
	13	18	90.0	86	4	US-09-152-060-3	Sequence 3, Appli
	14	18	90.0	271	3	US-09-106-182-18	Sequence 18, Appl
	15	18	90.0	271	4	US-09-227-357-5	Sequence 5, Appli
	16	18	90.0	271	4	US-09-280-839-8	Sequence 8, Appli
	17	18	90.0	271	4	US-09-411-977-20	Sequence 20, Appl
	18	18	90.0	271	4	US-09-479-729B-25	Sequence 25, Appl
	19	18	90.0	271	4	US-09-257-179-5	Sequence 5, Appli
	20	18	90.0	271	4	US-09-149-476-5	Sequence 5, Appli
	21	18	90.0	271	4	US-09-288-143-5	Sequence 5, Appli
	22	18	90.0	271	4	US-09-487-792-27	Sequence 27, Appl
	23	18	90.0	271	4	US-09-152-060-5	Sequence 5, Appli
c	24	17.4	87.0	20	3	US-08-837-635-2	Sequence 2, Appli
	25	16.8	84.0	28473	4	US-08-961-527-83	Sequence 83, Appl
	26	16.4	82.0	22	2	US-08-683-743-24	Sequence 24, Appl
c	27	15.4	77.0	1410	3	US-08-068-392-1	Sequence 1, Appli
c	28	15.4	77.0	1410	4	US-08-396-988-1	Sequence 1, Appli
c	29	15.4	77.0	2200	1	US-08-272-255-21	Sequence 21, Appl
c	30	15.4	77.0	2200	5	PCT-US95-08565-21	Sequence 21, Appl
	31	15.2	76.0	24	2	US-08-632-575B-56	Sequence 56, Appl
c	32	15.2	76.0	317	4	US-09-732-199A-10	Sequence 10, Appl
c	33	15.2	76.0	321	3	US-09-080-855-25	Sequence 25, Appl
	34	15.2	76.0	1599	4	US-09-277-565-16	Sequence 16, Appl
	35	15.2	76.0	7577	4	US-08-961-527-46	Sequence 46, Appl
c	36	15.2	76.0	8898	4	US-08-961-527-69	Sequence 69, Appl
	37	15	75.0	19	1	US-08-411-020-38	Sequence 38, Appl
c	38	15	75.0	19	1	US-08-411-020-39	Sequence 39, Appl
	39	15	75.0	19	1	US-08-410-779B-106	Sequence 106, App
c	40	15	75.0	19	1	US-08-410-779B-107	Sequence 107, App
	41	15	75.0	19	5	PCT-US95-04477-106	Sequence 106, App
c	42	15	75.0	19	5	PCT-US95-04477-107	Sequence 107, App
c	43	14.8	74.0	648	4	US-09-228-986-26	Sequence 26, Appl
	44	14.4	72.0	1050	1	US-08-204-196A-2	Sequence 2, Appli
	45	14.4	72.0	1809	1	US-08-204-196A-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-08-394-191-12/c
; Sequence 12, Application US/08394191
; Patent No. 5616489
;  GENERAL INFORMATION:
;    APPLICANT:  LEVY, David E.
;    TITLE OF INVENTION:  DNA SEQUENCE WHICH BINDS
;    TITLE OF INVENTION:  TRANSCRIPTIONAL REGULATORY PROTEINS ACTIVATED IN RESPONSE T
;    TITLE OF INVENTION:  VARIOUS CYTOKINES AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Avenue
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/394,191
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/121,931
;      FILING DATE:  September 15, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LIVNAT, SHMUEL
;      REGISTRATION NUMBER:  33,949
;      REFERENCE/DOCKET NUMBER:  15661-20010.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-394-191-12

  Query Match           100.0%;  Score 20;  DB 1;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 0.086;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
```

```
                    ||||||||||||||||||||
Db         20 CTGATTTCCCCGAAATGATG 1


RESULT 2
US-08-458-364-12/c
; Sequence 12, Application US/08458364
; Patent No. 5648217
;  GENERAL INFORMATION:
;    APPLICANT:  LEVY, David E.
;    TITLE OF INVENTION:  DNA SEQUENCE WHICH BINDS
;    TITLE OF INVENTION:  TRANSCRIPTIONAL REGULATORY PROTEINS ACTIVATED IN RESPONSE TO
;    TITLE OF INVENTION:  VARIOUS CYTOKINES AND USES THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Avenue
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,364
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/394,191
;      FILING DATE:  24-FEB-1995
;      APPLICATION NUMBER:  08/121,931
;      FILING DATE:  September 15, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LIVNAT, SHMUEL
;      REGISTRATION NUMBER:  33,949
;      REFERENCE/DOCKET NUMBER:  15661-20010.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-458-364-12

  Query Match             100.0%;  Score 20;  DB 1;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 0.086;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              ||||||||||||||||||||
Db         20 CTGATTTCCCCGAAATGATG 1


RESULT 3
US-08-837-635-1
; Sequence 1, Application US/08837635
; Patent No. 6007998
;  GENERAL INFORMATION:
;    APPLICANT:  ROSENBLUM, CHARLES, I.
;    APPLICANT:  VAN DER PLOEG, LEONARDUS, H.T.
;    APPLICANT:  QURESHI, SAJJAD, A.
;    APPLICANT:  CULLY, DORIS, F.
;    APPLICANT:  HESS, JOHN W.
;    APPLICANT:  TOTA, MICHAEL, R.
;    APPLICANT:  CHEN, FANG
;    TITLE OF INVENTION:  LEPTIN ASSAY
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JOANNE M. GIESSER - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/837,635
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/016,051
;      FILING DATE:  22-APR-1996
;      APPLICATION NUMBER:  60/031,002
;      FILING DATE:  15-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GIESSER, JOANNE M
;      REGISTRATION NUMBER:  32,838
;      REFERENCE/DOCKET NUMBER:  19686Y
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-3046
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other
US-08-837-635-1

  Query Match             92.0%;  Score 18.4;  DB 3;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 0.58;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATTTCCCCGAAATGATG 20
              |||||||||||||||||| |
Db          1 CTGATTTCCCCGAAATGACG 20


RESULT 4
US-09-106-182-16
; Sequence 16, Application US/09106182
; Patent No. 6046035
;  GENERAL INFORMATION:
;    APPLICANT:  Shi, Yanggu
;    APPLICANT:  Ruben, Steve
;    TITLE OF INVENTION:  Cardiotrophin-Like Cytokine
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc
;      STREET:  9410 Key West Ave
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/106,182
;      FILING DATE:  Herewith
;      CLASSIFICATION:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/051,053
;      FILING DATE:  30-JUN-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF385
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  86 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-106-182-16

  Query Match             90.0%;  Score 18;  DB 3;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57


RESULT 5
US-09-227-357-3
; Sequence 3, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
;  APPLICANT: Fischer et al.
;  TITLE OF INVENTION: 123 Human Secreted Proteins
;  FILE REFERENCE: PZ010P1
;  CURRENT APPLICATION NUMBER: US/09/227,357
;  CURRENT FILING DATE: 1999-01-08
;  EARLIER APPLICATION NUMBER: PCT/US98/13684
;  EARLIER FILING DATE: 1998-07-07
;  EARLIER APPLICATION NUMBER: 60/051,926
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,793
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,925
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,929
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,803
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,732
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,931
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,932
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,916
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,930
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,918
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,920
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,733
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,795
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,919
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,928
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/055,722
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,723
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,948
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,949
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,953
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,950
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,947
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,964
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/056,360
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,684
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,984
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,954
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/058,785
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,664
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,660
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,661
;  EARLIER FILING DATE: 1997-09-12
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 86
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-227-357-3

  Query Match             90.0%;  Score 18;  DB 4;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57


RESULT 6
US-09-280-839-6
; Sequence 6, Application US/09280839
; Patent No. 6365369
; GENERAL INFORMATION:
;  APPLICANT: Endress, Gregory A.
;  APPLICANT: Rosen, Craig A.
;  TITLE OF INVENTION: Prostate Specific Secreted Protein
;  FILE REFERENCE: PF457
;  CURRENT APPLICATION NUMBER: US/09/280,839
;  CURRENT FILING DATE: 1999-03-30
;  EARLIER APPLICATION NUMBER: 60/080,311
;  EARLIER FILING DATE: 1998-04-01
;  EARLIER APPLICATION NUMBER: 60/080,898
;  EARLIER FILING DATE: 1998-04-07
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 86
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-280-839-6

  Query Match             90.0%;  Score 18;  DB 4;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
```

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57
```

RESULT 7
US-09-411-977-18
; Sequence 18, Application US/09411977
; Patent No. 6372473
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease
; FILE REFERENCE: PF378P1
; CURRENT APPLICATION NUMBER: US/09/411,977
; CURRENT FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: 09/084,491
; EARLIER FILING DATE: 1998-05-27
; EARLIER APPLICATION NUMBER: 60/048,000
; EARLIER FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 86
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-411-977-18

Query Match 90.0%; Score 18; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57
```

RESULT 8
US-09-479-729B-23
; Sequence 23, Application US/09479729B
; Patent No. 6391589
; GENERAL INFORMATION:
; APPLICANT: Olsen, et al
; TITLE OF INVENTION: Human Chemokine Beta-10 Mutant Polypeptides
; FILE REFERENCE: PF504
; CURRENT APPLICATION NUMBER: US/09/479,729B
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: PCT/US94/09484
; PRIOR FILING DATE: 1994-08-23
; PRIOR APPLICATION NUMBER: 08/458,355
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: 08/462,967
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: 60/115,439
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 86
; TYPE: DNA
; ORGANISM: oligonucleotide
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: (1)..(86)
; OTHER INFORMATION: 5' primer to generate GAS-SV40 construct.
US-09-479-729B-23

Query Match 90.0%; Score 18; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57
```

RESULT 9
US-09-257-179-3
; Sequence 3, Application US/09257179
; Patent No. 6410709
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 29 Human Secreted Proteins
; FILE REFERENCE: PZ015P1
; CURRENT APPLICATION NUMBER: US/09/257,179
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: PCT/US98/17709
; EARLIER FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/056,270
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,271
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,247
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,073
; EARLIER FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 86
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-257-179-3

Query Match 90.0%; Score 18; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57
```

RESULT 10
US-09-149-476-3
; Sequence 3, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22

```
;  EARLIER APPLICATION NUMBER: 60/056,876
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,881
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,909
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,875
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,862
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,887
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,908
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/057,650
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: 60/056,884
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/057,669
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: 60/049,610
;  EARLIER FILING DATE: 1997-06-13
;  EARLIER APPLICATION NUMBER: 60/061,060
;  EARLIER FILING DATE: 1997-10-02

  Query Match             90.0%;  Score 18;  DB 4;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57


RESULT 11
US-09-288-143-3
; Sequence 3, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
;  APPLICANT: Brewer et al.
;  TITLE OF INVENTION: 53 Human Secreted Proteins
;  FILE REFERENCE: PZ018P1
;  CURRENT APPLICATION NUMBER: US/09/288,143
;  CURRENT FILING DATE: 1999-04-08
;  EARLIER APPLICATION NUMBER: PCT/US98/21142
;  EARLIER FILING DATE: 1998-10-08
;  EARLIER APPLICATION NUMBER: 60/061,463
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,529
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/071,498
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,527
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,536
;  EARLIER FILING DATE: 1997-10-09
;  EARLIER APPLICATION NUMBER: 60/061,532
;  EARLIER FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 219
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 86
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-288-143-3

  Query Match             90.0%;  Score 18;  DB 4;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57


RESULT 12
US-09-487-792-25
; Sequence 25, Application US/09487792
; Patent No. 6433145
; GENERAL INFORMATION:
;  APPLICANT: Human Genome Sciences, Inc.
;  TITLE OF INVENTION: Keratinocyte Derived Interferon
;  FILE REFERENCE: PF482P1
;  CURRENT APPLICATION NUMBER: US/09/487,792
;  CURRENT FILING DATE: 2000-01-20
;  EARLIER APPLICATION NUMBER: 60/093,643
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: PCT/US99/16424
;  EARLIER FILING DATE: 1999-07-21
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 86
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-487-792-25

  Query Match             90.0%;  Score 18;  DB 4;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57


RESULT 13
US-09-152-060-3
; Sequence 3, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 28 Human Secreted Proteins
;  FILE REFERENCE: PZ003P1.US
;  CURRENT APPLICATION NUMBER: US/09/152,060
;  CURRENT FILING DATE: 1998-09-11
;  EARLIER APPLICATION NUMBER: PCT/US98/04858
;  EARLIER FILING DATE: 1998-03-12
;  EARLIER APPLICATION NUMBER: 60/040,762
;  EARLIER FILING DATE: 1997-03-14
;  EARLIER APPLICATION NUMBER: 60/040,710
;  EARLIER FILING DATE: 1997-03-14
;  EARLIER APPLICATION NUMBER: 60/050,934
;  EARLIER FILING DATE: 1997-05-30
;  EARLIER APPLICATION NUMBER: 60/048,100
;  EARLIER FILING DATE: 1997-05-30
;  EARLIER APPLICATION NUMBER: 60/048,357
;  EARLIER FILING DATE: 1997-05-30
;  EARLIER APPLICATION NUMBER: 60/048,189
;  EARLIER FILING DATE: 1997-05-30
;  EARLIER APPLICATION NUMBER: 60/057,765
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: 60/048,970
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/068,368
;  EARLIER FILING DATE: 1997-12-19
;  NUMBER OF SEQ ID NOS: 118
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 86
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-152-060-3
```

```
  Query Match             90.0%;  Score 18;  DB 4;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         40 TGATTTCCCCGAAATGAT 57


RESULT 14
US-09-106-182-18
; Sequence 18, Application US/09106182
; Patent No. 6046035
;  GENERAL INFORMATION:
;    APPLICANT:  Shi, Yanggu
;    APPLICANT:  Ruben, Steve
;    TITLE OF INVENTION:  Cardiotrophin-Like Cytokine
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc
;      STREET:  9410 Key West Ave
;      CITY:  Rockville
;      STATE:  MD
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/106,182
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/051,053
;      FILING DATE:  30-JUN-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PF385
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-309-8504
;      TELEFAX:  301-309-8439
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  271 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-106-182-18

  Query Match             90.0%;  Score 18;  DB 3;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         36 TGATTTCCCCGAAATGAT 53


RESULT 15
US-09-227-357-5
; Sequence 5, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
;  APPLICANT: Fischer et al.
;  TITLE OF INVENTION: 123 Human Secreted Proteins
;  FILE REFERENCE: PZ010P1
;  CURRENT APPLICATION NUMBER: US/09/227,357
;  CURRENT FILING DATE: 1999-01-08
;  EARLIER APPLICATION NUMBER: PCT/US98/13684
;  EARLIER FILING DATE: 1998-07-07
;  EARLIER APPLICATION NUMBER: 60/051,926
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,793
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,925
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,929
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,803
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,732
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,931
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,932
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,916
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,930
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,918
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,920
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,733
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/052,795
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,919
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/051,928
;  EARLIER FILING DATE: 1997-07-08
;  EARLIER APPLICATION NUMBER: 60/055,722
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,723
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,948
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,949
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,953
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,950
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,947
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,964
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/056,360
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,684
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,984
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/055,954
;  EARLIER FILING DATE: 1997-08-18
;  EARLIER APPLICATION NUMBER: 60/058,785
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,664
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,660
;  EARLIER FILING DATE: 1997-09-12
;  EARLIER APPLICATION NUMBER: 60/058,661
;  EARLIER FILING DATE: 1997-09-12
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 271
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-227-357-5
```

Query Match 90.0%; Score 18; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 TGATTTCCCCGAAATGAT 19
              ||||||||||||||||||
Db         36 TGATTTCCCCGAAATGAT 53
```

Search completed: June 26, 2003, 16:21:12
Job time : 32.2888 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 10:58:25 ; Search time 30.2174 Seconds
(without alignments)
202.980 Million cell updates/sec

Title: US-09-355-254F-21
Perfect score: 20
Sequence: 1 gtatttcccagaaaaggaac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	1	US-08-276-099A-3	Sequence 3, Appli
2	20	100.0	20	1	US-08-393-333-3	Sequence 3, Appli
3	20	100.0	20	1	US-08-408-318-3	Sequence 3, Appli
4	20	100.0	20	1	US-08-781-890-3	Sequence 3, Appli
5	20	100.0	20	1	US-08-839-164-3	Sequence 3, Appli
6	20	100.0	20	4	US-09-511-625B-9	Sequence 9, Appli
7	20	100.0	20	4	US-09-511-625B-14	Sequence 14, Appl
8	20	100.0	30	4	US-08-956-653A-25	Sequence 25, Appl
9	20	100.0	100	4	US-09-522-217-59	Sequence 59, Appl
c 10	20	100.0	100	4	US-09-522-217-60	Sequence 60, Appl
11	18.4	92.0	62	4	US-09-003-903-5	Sequence 5, Appli
12	17.4	87.0	26	4	US-09-003-903-1	Sequence 1, Appli
13	17	85.0	17	1	US-08-369-796-22	Sequence 22, Appl
14	17	85.0	17	2	US-08-852-091-22	Sequence 22, Appl
c 15	17	85.0	17	4	US-09-178-973B-14	Sequence 14, Appl
c 16	17	85.0	17	4	US-09-419-568F-14	Sequence 14, Appl
c 17	17	85.0	17	4	US-09-354-243B-14	Sequence 14, Appl
18	17	85.0	17	5	PCT-US95-17025-22	Sequence 22, Appl
19	17	85.0	49	4	US-08-956-653A-24	Sequence 24, Appl
20	16.8	84.0	1001	4	US-09-641-638-104	Sequence 104, App
c 21	16.4	82.0	36651	4	US-09-738-894A-3	Sequence 3, Appli
22	16	80.0	25	2	US-08-820-754-17	Sequence 17, Appl
23	16	80.0	25	3	US-08-956-652-17	Sequence 17, Appl
24	16	80.0	25	3	US-08-956-869-17	Sequence 17, Appl
25	16	80.0	25	3	US-08-948-547-17	Sequence 17, Appl
26	16	80.0	25	4	US-08-956-653A-17	Sequence 17, Appl
27	16	80.0	30	1	US-08-386-728-4	Sequence 4, Appli
28	16	80.0	30	5	PCT-US96-01768-4	Sequence 4, Appli
c 29	15.8	79.0	1364	2	US-08-872-302-3	Sequence 3, Appli
30	15.8	79.0	1383	1	US-08-289-709-2	Sequence 2, Appli
31	15.8	79.0	1383	1	US-08-602-656-2	Sequence 2, Appli
c 32	15.8	79.0	98844	4	US-09-791-211-10	Sequence 10, Appl
c 33	15.4	77.0	1600	4	US-07-861-458C-37	Sequence 37, Appl
c 34	15.4	77.0	1607	6	5196333-3	Patent No. 5196333
c 35	15.4	77.0	10614	1	US-08-135-511-35	Sequence 35, Appl
c 36	15.4	77.0	10614	1	US-08-187-453-35	Sequence 35, Appl
37	15.2	76.0	243	4	US-09-134-001C-1717	Sequence 1717, Ap
38	15.2	76.0	8789	1	US-08-328-254-5	Sequence 5, Appli
39	15.2	76.0	10136	1	US-08-353-700-2	Sequence 2, Appli
40	15.2	76.0	10136	5	PCT-US95-16216-2	Sequence 2, Appli
c 41	15	75.0	15	1	US-08-141-499A-13	Sequence 13, Appl
c 42	15	75.0	15	1	US-08-467-940-13	Sequence 13, Appl
c 43	15	75.0	15	1	US-08-633-772-13	Sequence 13, Appl
44	15	75.0	1266	4	US-09-134-078-3	Sequence 3, Appli
45	15	75.0	7577	4	US-08-961-527-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-08-276-099A-3
; Sequence 3, Application US/08276099A
; Patent No. 5591825
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L
; APPLICANT: Hou, Jinzhao
; TITLE OF INVENTION: INTERLEUKIN-4 SIGNAL TRANSDUCERS AND
; TITLE OF INVENTION: BINDING ASSAYS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,099A
; FILING DATE: 15-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59451-1/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-276-099A-3

Query Match 100.0%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 1 GTATTTCCCAGAAAAGGAAC 20

```
RESULT 2
US-08-393-333-3
; Sequence 3, Application US/08393333
; Patent No. 5618693
;  GENERAL INFORMATION:
;    APPLICANT:  McKnight, Steven L
;    APPLICANT:  Hou, Jinzhao
;    APPLICANT:  Schindler, Ulrike
;    TITLE OF INVENTION:  INTERLEUKIN-2 SIGNAL TRANSDUCERS AND
;    TITLE OF INVENTION:  BINDING ASSAYS
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/393,333
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-60778/RAO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-393-333-3

  Query Match             100.0%;  Score 20;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 3
US-08-408-318-3
; Sequence 3, Application US/08408318
; Patent No. 5639858
;  GENERAL INFORMATION:
;    APPLICANT:  Hoey, Timothy
;    TITLE OF INVENTION:  Human Signal Transducers and Binding
;    TITLE OF INVENTION:  Assays
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herber
;      STREET:  850 Hansen Way, #200
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/408,318
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-60845
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-494-8700
;      TELEFAX:  415-494-8771
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-408-318-3

  Query Match             100.0%;  Score 20;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 4
US-08-781-890-3
; Sequence 3, Application US/08781890
; Patent No. 5710266
;  GENERAL INFORMATION:
;    APPLICANT:  McKnight, Steven L
;    APPLICANT:  Hou, Jinzhao
;    TITLE OF INVENTION:  INTERLEUKIN-4 SIGNAL TRANSDUCERS AND
;    TITLE OF INVENTION:  BINDING ASSAYS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,890
;      FILING DATE:  05-JAN-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/276,099
;      FILING DATE:  15-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard Aron
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-59451-1/RAO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-781-890-3

  Query Match            100.0%;  Score 20;  DB 1;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 0.17;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 5
US-08-839-164-3
; Sequence 3, Application US/08839164
; Patent No. 5756700
;  GENERAL INFORMATION:
;    APPLICANT:  Hoey, Timothy
;    TITLE OF INVENTION:  Human Signal Transducers and Binding
;    TITLE OF INVENTION:  Assays
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herber
;      STREET:  850 Hansen Way, #200
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/839,164
;      FILING DATE:  23-APR-1997
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/408,318
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Osman, Richard A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  A-60845
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-494-8700
;      TELEFAX:  415-494-8771
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-839-164-3

  Query Match            100.0%;  Score 20;  DB 1;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 0.17;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 6
US-09-511-625B-9
; Sequence 9, Application US/09511625B
; Patent No. 6368828
; GENERAL INFORMATION:
;  APPLICANT:  LaRochelle, William J.
;  APPLICANT:  Patel, Bharvin
;  APPLICANT:  Pierce, Jacalyn H.
;  TITLE OF INVENTION:  ATTENUATED AND DOMINANT NEGATIVEVARIANT
;  TITLE OF INVENTION:  cDNAs OF STAT6: STAT6b AND STAT6c
;  FILE REFERENCE:  14014.0300u1
;  CURRENT APPLICATION NUMBER: US/09/511,625B
;  CURRENT FILING DATE:  2000-02-23
;  PRIOR APPLICATION NUMBER:  PCT/US98/17821
;  PRIOR FILING DATE:  1998-08-27
;  PRIOR APPLICATION NUMBER:  60/070,397
;  PRIOR FILING DATE:  1998-01-05
;  PRIOR APPLICATION NUMBER:  60/056,075
;  PRIOR FILING DATE:  1997-08-27
;  NUMBER OF SEQ ID NOS:  70
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial Sequence:/No. 6368828e -
;   OTHER INFORMATION: synthetic construct
US-09-511-625B-9

  Query Match            100.0%;  Score 20;  DB 4;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 0.17;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 7
US-09-511-625B-14
; Sequence 14, Application US/09511625B
; Patent No. 6368828
; GENERAL INFORMATION:
;  APPLICANT:  LaRochelle, William J.
;  APPLICANT:  Patel, Bharvin
;  APPLICANT:  Pierce, Jacalyn H.
;  TITLE OF INVENTION:  ATTENUATED AND DOMINANT NEGATIVEVARIANT
;  TITLE OF INVENTION:  cDNAs OF STAT6: STAT6b AND STAT6c
;  FILE REFERENCE:  14014.0300u1
;  CURRENT APPLICATION NUMBER: US/09/511,625B
;  CURRENT FILING DATE:  2000-02-23
;  PRIOR APPLICATION NUMBER:  PCT/US98/17821
;  PRIOR FILING DATE:  1998-08-27
;  PRIOR APPLICATION NUMBER:  60/070,397
;  PRIOR FILING DATE:  1998-01-05
;  PRIOR APPLICATION NUMBER:  60/056,075
;  PRIOR FILING DATE:  1997-08-27
;  NUMBER OF SEQ ID NOS:  70
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial Sequence:/No. 6368828e -
;   OTHER INFORMATION: synthetic construct
US-09-511-625B-14

  Query Match            100.0%;  Score 20;  DB 4;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 0.17;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 8
```

```
US-08-956-653A-25
; Sequence 25, Application US/08956653A
; Patent No. 6338949
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,653A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-195
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;      DESCRIPTION:  oligonucleotide probe Fc RI.2
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-956-653A-25

  Query Match           100.0%;  Score 20;  DB 4;  Length 30;
  Best Local Similarity  100.0%;  Pred. No. 0.18;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          6 GTATTTCCCAGAAAAGGAAC 25


RESULT 9
US-09-522-217-59
; Sequence 59, Application US/09522217
; Patent No. 6307024
; GENERAL INFORMATION:
;  APPLICANT: No. 6307024ak, Julia E.
;  APPLICANT: Presnell, Scott R.
;  APPLICANT: Sprecher, Cindy A.
;  APPLICANT: Foster, Donald C.
;  APPLICANT: Holly, Richard D.
;  APPLICANT: Gross, Jane A.
;  APPLICANT: Johnston, Janet V.
;  APPLICANT: Nelson, Andrew J.
;  APPLICANT: Dillon, Stacey R.
;  APPLICANT: Hammond, Angela K.
;  TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
;  FILE REFERENCE: 99-16
;  CURRENT APPLICATION NUMBER: US/09/522,217
;  CURRENT FILING DATE: 2000-03-09
;  EARLIER APPLICATION NUMBER: US 60/123,547
;  EARLIER FILING DATE: 1999-03-09
;  EARLIER APPLICATION NUMBER: US 60/123,904
;  EARLIER FILING DATE: 1999-03-11
;  EARLIER APPLICATION NUMBER: US 60/142,013
;  EARLIER FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 59
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide ZC12749
US-09-522-217-59

  Query Match           100.0%;  Score 20;  DB 4;  Length 100;
  Best Local Similarity  100.0%;  Pred. No. 0.23;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db         44 GTATTTCCCAGAAAAGGAAC 63


RESULT 10
US-09-522-217-60/c
; Sequence 60, Application US/09522217
; Patent No. 6307024
; GENERAL INFORMATION:
;  APPLICANT: No. 6307024ak, Julia E.
;  APPLICANT: Presnell, Scott R.
;  APPLICANT: Sprecher, Cindy A.
;  APPLICANT: Foster, Donald C.
;  APPLICANT: Holly, Richard D.
;  APPLICANT: Gross, Jane A.
;  APPLICANT: Johnston, Janet V.
;  APPLICANT: Nelson, Andrew J.
;  APPLICANT: Dillon, Stacey R.
;  APPLICANT: Hammond, Angela K.
;  TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
;  FILE REFERENCE: 99-16
;  CURRENT APPLICATION NUMBER: US/09/522,217
;  CURRENT FILING DATE: 2000-03-09
;  EARLIER APPLICATION NUMBER: US 60/123,547
;  EARLIER FILING DATE: 1999-03-09
;  EARLIER APPLICATION NUMBER: US 60/123,904
;  EARLIER FILING DATE: 1999-03-11
;  EARLIER APPLICATION NUMBER: US 60/142,013
;  EARLIER FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 60
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
```

```
;  FEATURE:
;   OTHER INFORMATION: Oligonucleotide ZC12748
US-09-522-217-60

  Query Match             100.0%;  Score 20;  DB 4;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db         61 GTATTTCCCAGAAAAGGAAC 42


RESULT 11
US-09-003-903-5
; Sequence 5, Application US/09003903
; Patent No. 6265160
; GENERAL INFORMATION:
;  APPLICANT: The Government of the United States of America
;  APPLICANT: as Represented by the Secretary, Department of
;  APPLICANT: Health and Human Services
;  APPLICANT: Leonard, Warren J.
;  TITLE OF INVENTION: METHOD OF IDENTIFYING INHIBITORS OF THE
;  TITLE OF INVENTION: JAK-STAT PATHWAY
;  FILE REFERENCE: NIH118.001C1
;  CURRENT APPLICATION NUMBER: US/09/003,903
;  CURRENT FILING DATE: 1998-01-07
;  EARLIER APPLICATION NUMBER: PCT/US96/11206
;  EARLIER FILING DATE: 1996-07-02
;  EARLIER APPLICATION NUMBER: 60/000,971
;  EARLIER FILING DATE: 1995-07-07
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 62
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-003-903-5

  Query Match             92.0%;  Score 18.4;  DB 4;  Length 62;
  Best Local Similarity   95.0%;  Pred. No. 1.3;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              |||||||||||||||||| |
Db          5 GTATTTCCCAGAAAAGGATC 24


RESULT 12
US-09-003-903-1
; Sequence 1, Application US/09003903
; Patent No. 6265160
; GENERAL INFORMATION:
;  APPLICANT: The Government of the United States of America
;  APPLICANT: as Represented by the Secretary, Department of
;  APPLICANT: Health and Human Services
;  APPLICANT: Leonard, Warren J.
;  TITLE OF INVENTION: METHOD OF IDENTIFYING INHIBITORS OF THE
;  TITLE OF INVENTION: JAK-STAT PATHWAY
;  FILE REFERENCE: NIH118.001C1
;  CURRENT APPLICATION NUMBER: US/09/003,903
;  CURRENT FILING DATE: 1998-01-07
;  EARLIER APPLICATION NUMBER: PCT/US96/11206
;  EARLIER FILING DATE: 1996-07-02
;  EARLIER APPLICATION NUMBER: 60/000,971
;  EARLIER FILING DATE: 1995-07-07
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 26
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-003-903-1

  Query Match             87.0%;  Score 17.4;  DB 4;  Length 26;
  Best Local Similarity   94.7%;  Pred. No. 3.5;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAA 19
              ||||||||||||||||| |
Db          6 GTATTTCCCAGAAAAGGGA 24


RESULT 13
US-08-369-796-22
; Sequence 22, Application US/08369796
; Patent No. 5716622
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA synthetic probe
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-369-796-22

  Query Match             85.0%;  Score 17;  DB 1;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGG 17
              |||||||||||||||||
Db          1 GTATTTCCCAGAAAAGG 17


RESULT 14
US-08-852-091-22
; Sequence 22, Application US/08852091
```

```
; Patent No. 5883228
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,091
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA synthetic probe
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-852-091-22

  Query Match             85.0%;  Score 17;  DB 2;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGG 17
              |||||||||||||||||
Db          1 GTATTTCCCAGAAAAGG 17


RESULT 15
US-09-178-973B-14/c
; Sequence 14, Application US/09178973B
; Patent No. 6274710
; GENERAL INFORMATION:
;  APPLICANT: Dumoutier, Laure
;  APPLICANT:  Louhed, Jamila
;  APPLICANT:  Renauld, Jean-Christophe
;  TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac
;  TITLE OF INVENTION:  (TIFs)
;  TITLE OF INVENTION:  The Proteins Encoded, and Uses Thereof
;  FILE REFERENCE: LUD 5543
;  CURRENT APPLICATION NUMBER: US/09/178,973B
;  CURRENT FILING DATE:  1998-10-26
;  NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 14
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-178-973B-14

  Query Match             85.0%;  Score 17;  DB 4;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 5.1;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGG 17
              |||||||||||||||||
Db         17 GTATTTCCCAGAAAAGG 1


Search completed: June 26, 2003, 16:21:14
Job time : 31.2888 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 11:19:15 ; Search time 1529.13 Seconds
(without alignments)
211.826 Million cell updates/sec

Title: US-09-355-254F-8
Perfect score: 20
Sequence: 1 gattgcctgacgtcagagag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	17.4	87.0	509	17	AZ456210	AZ456210 1M0258J21
c	2	17.4	87.0	525	17	AZ930991	AZ930991 474.dhz60
c	3	16.8	84.0	580	12	BF571725	BF571725 602076288
c	4	16.8	84.0	581	12	BG529082	BG529082 602579184
c	5	16.8	84.0	812	12	BE865809	BE865809 601678202
c	6	16.8	84.0	818	12	BG121636	BG121636 602351594
	7	16.8	84.0	984	17	AG085708	AG085708 Pan trogl
c	8	16.4	82.0	421	9	AI446142	AI446142 tj07d08.x
c	9	16.4	82.0	455	9	AI377721	AI377721 te56e04.x
c	10	16.4	82.0	462	9	AU147973	AU147973 AU147973
c	11	16.4	82.0	468	9	AA524774	AA524774 nh33f04.s
c	12	16.4	82.0	476	12	BF439459	BF439459 nab64g11.
c	13	16.4	82.0	570	17	AQ617221	AQ617221 HS_5152_B
	14	16.4	82.0	618	17	AQ392594	AQ392594 CITBI-E1-
c	15	16.4	82.0	630	17	AQ884463	AQ884463 HS_5510_A
	16	16.4	82.0	722	10	AV764299	AV764299 AV764299
	17	16.4	82.0	737	12	BG707030	BG707030 602670345
	18	16.4	82.0	1094	17	CNS05K18	AL340901 Tetraodon
	19	16	80.0	473	10	BB697276	BB697276 BB697276
	20	16	80.0	668	17	BH306678	BH306678 CH230-98F
c	21	16	80.0	1032	17	CNS01VJ6	AL169179 Tetraodon
	22	15.8	79.0	123	9	AA750965	AA750965 ISKP0091
	23	15.8	79.0	281	9	AA795454	AA795454 vv20d03.r
	24	15.8	79.0	292	10	BB306386	BB306386 BB306386
c	25	15.8	79.0	304	10	BE191014	BE191014 sn83f10.y
c	26	15.8	79.0	383	14	H64068	H64068 yr58f04.r1
	27	15.8	79.0	410	12	BF709136	BF709136 MI-P-AY0-
	28	15.8	79.0	424	10	AV407792	AV407792 AV407792
	29	15.8	79.0	435	17	B41297	B41297 HS-1053-B1-
c	30	15.8	79.0	442	10	AW211525	AW211525 uo82b04.y
c	31	15.8	79.0	444	13	BM174442	BM174442 Tm_ad_29E
	32	15.8	79.0	452	17	AQ179770	AQ179770 HS_3185_B
c	33	15.8	79.0	484	17	AQ460036	AQ460036 HS_5136_A
	34	15.8	79.0	488	14	BQ312839	BQ312839 RC0-BN028
	35	15.8	79.0	510	17	AQ809522	AQ809522 HS_4646_B
	36	15.8	79.0	528	17	AQ471562	AQ471562 CITBI-E1-
	37	15.8	79.0	529	17	BH200967	BH200967 Sm1-43K17
	38	15.8	79.0	533	10	AW671751	AW671751 LG1_351_E
c	39	15.8	79.0	543	12	BE862853	BE862853 UI-M-BG0-
c	40	15.8	79.0	553	17	AQ672151	AQ672151 HS_2160_B
	41	15.8	79.0	596	12	BF374971	BF374971 CM2-ST018
	42	15.8	79.0	605	17	AQ532396	AQ532396 RPCI-11-3
c	43	15.8	79.0	612	10	AW841675	AW841675 RC3-CN001
c	44	15.8	79.0	612	17	AZ083796	AZ083796 RPCI-23-2
	45	15.8	79.0	613	17	AQ305574	AQ305574 HS_2034_A

ALIGNMENTS

RESULT 1
AZ456210
LOCUS AZ456210 509 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0258J21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0258J21 R, DNA sequence.
ACCESSION AZ456210
VERSION AZ456210.1 GI:10614335
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 509)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

```
           Plate: 0258  row: J  column: 21
           Seq primer: CACACAGGAAACAGCTATGACC
           Class: plasmid ends
           High quality sequence stop: 509.
FEATURES             Location/Qualifiers
     source          1. .509
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0258J21"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT      176 a     93 c    132 g    108 t
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 17;  Length 509;
  Best Local Similarity   94.7%;  Pred. No. 8.7e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 ATTGCCTGACGTCAGAGAG 20
              ||||||||| |||||||||
Db        226 ATTGCCTGATGTCAGAGAG 244


RESULT 2
AZ930991/c
LOCUS       AZ930991                525 bp    DNA     linear   GSS 01-APR-2001
DEFINITION  474.dhz60h07.s1 Saccharomyces unisporus NRRL Y-1556 Saccharomyces
            unisporus genomic clone 474.dhz60h07.s1, DNA sequence.
ACCESSION   AZ930991
VERSION     AZ930991.1  GI:13501901
KEYWORDS    GSS.
SOURCE      Saccharomyces unisporus.
  ORGANISM  Saccharomyces unisporus
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 525)
  AUTHORS   Cliften,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish
            ,W.R., Waterston,R.H. and Johnston,M.
  TITLE     Surveying Saccharomyces genomes to identify functional elements by
            comparative DNA sequence analysis
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Johnston M
            Department of Genetics
            Washington University Medical School
            Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
            Tel: 314 362 2735
            Fax: 314 362 7855
            Email: mj@genetics.wustl.edu
            Class: random plasmid subclone.
FEATURES             Location/Qualifiers
     source          1. .525
                     /organism="Saccharomyces unisporus"
                     /strain="NRRL Y-1556 (CBS 398)"
                     /db_xref="taxon:27294"
                     /clone="474.dhz60h07.s1"
                     /clone_lib="Saccharomyces unisporus NRRL Y-1556"
                     /note="Random genomic sequence "
BASE COUNT      133 a     119 c     77 g     196 t
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 17;  Length 525;
  Best Local Similarity   94.7%;  Pred. No. 8.8e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGA 19
              |||||||||||||||||| |
Db        384 GATTGCCTGACGTCAGAAA 366


RESULT 3
BF571725/c
LOCUS       BF571725                580 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602076288F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4243820 5',
            mRNA sequence.
ACCESSION   BF571725
VERSION     BF571725.1  GI:11645437
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1054  row: o  column: 21
            High quality sequence start: 6
            High quality sequence stop: 488.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4243820"
                     /clone_lib="NIH_MGC_62"
                     /tissue_type="melanotic melanoma, high MDR"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.75 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      131 a     139 c     150 g     160 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 12;  Length 580;
  Best Local Similarity   90.0%;  Pred. No. 1.7e+03;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              |||||||||| ||||||||
```

```
Db        172 GATTGCCTGAGCTCAGAGAG 153


RESULT 4
BG529082/c
LOCUS       BG529082                581 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602579184F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4713261 5',
            mRNA sequence.
ACCESSION   BG529082
VERSION     BG529082.1  GI:13520619
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1556  row: g  column: 22
            High quality sequence stop: 566.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4713261"
                     /clone_lib="NIH_MGC_60"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); Double-stranded cDNA was prepared from cell line RNA.
                     5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).  Average
                     insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
                     contained inserts  by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library."
BASE COUNT      175 a    110 c    100 g    196 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 12;  Length 581;
  Best Local Similarity   90.0%;  Pred. No. 1.7e+03;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              |||||||||| | |||||||
Db         56 GATTGCCTGAAGGCAGAGAG 37


RESULT 5
BE865809/c
LOCUS       BE865809                812 bp    mRNA    linear   EST 27-SEP-2000
DEFINITION  601678202F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3960888 5',
            mRNA sequence.
ACCESSION   BE865809
VERSION     BE865809.1  GI:10314585
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 812)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM844  row: c  column: 01
            High quality sequence stop: 478.
FEATURES             Location/Qualifiers
     source          1. .812
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3960888"
                     /clone_lib="NIH_MGC_53"
                     /tissue_type="carcinoma, cell line"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: bladder; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); Double-stranded cDNA was prepared from cell line RNA.
                     5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C, G, or T).  Average
                     insert size 1.55 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      211 a    219 c    161 g    221 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 12;  Length 812;
  Best Local Similarity   90.0%;  Pred. No. 2e+03;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              |||||||||| | |||||||
Db        381 GATTGCCTGAAGGCAGAGAG 362


RESULT 6
BG121636/c
LOCUS       BG121636                818 bp    mRNA    linear   EST 30-JAN-2001
DEFINITION  602351594F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4449935 5',
            mRNA sequence.
ACCESSION   BG121636
VERSION     BG121636.1  GI:12615145
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 818)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10234  row: k  column: 24
            High quality sequence stop: 559.
```

```
FEATURES             Location/Qualifiers
     source          1. .818
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4449935"
                     /clone_lib="NIH_MGC_90"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.7 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      211 a    192 c    192 g    223 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 12;  Length 818;
  Best Local Similarity   90.0%;  Pred. No. 2e+03;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              |||||||||| | |||||||
Db        528 GATTGCCTGAAGGCAGAGAG 509


RESULT 7
AG085708
LOCUS       AG085708                984 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-083N10.R, genomic survey sequence.
ACCESSION   AG085708
VERSION     AG085708.1  GI:16637510
KEYWORDS    GSS.
SOURCE      Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
            BAC Library clone:PTB-083N10.R.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 984)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                 Sequencing: M13Rev
             LIBRARY
                 Vector    : pKS145
                 R.Site 1  : SacI
                 R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .984
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="PTB-083N10.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT      292 a    249 c    192 g    250 t      1 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 17;  Length 984;
  Best Local Similarity   90.0%;  Pred. No. 2.2e+03;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              || |||||||||||| |||
Db        243 GACTGCCTGACGTCAGCGAG 262


RESULT 8
AI446142/c
LOCUS       AI446142                421 bp    mRNA    linear   EST 09-MAR-1999
DEFINITION  tj07d08.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2140815 3'
            similar to contains Alu repetitive element;contains element MER22
            repetitive element ;, mRNA sequence.
ACCESSION   AI446142
VERSION     AI446142.1  GI:4293138
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 404.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2140815"
                     /clone_lib="NCI_CGAP_Gas4"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
BASE COUNT       81 a    120 c    119 g     99 t      2 others
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 9;  Length 421;
  Best Local Similarity   94.4%;  Pred. No. 2.3e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAG 18
              |||||||||| |||||||
Db        227 GATTGCCTGAGGTCAGAG 210


RESULT 9
AI377721/c
LOCUS       AI377721                455 bp    mRNA    linear   EST 28-MAR-1999
DEFINITION  te56e04.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2090718 3', mRNA sequence.
ACCESSION   AI377721
VERSION     AI377721.1  GI:4187574
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 700   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .455
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2090718"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT      130 a     92 c     97 g    136 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 9;  Length 455;
  Best Local Similarity   94.4%;  Pred. No. 2.4e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAG 18
              |||||||||| ||||||
Db        258 GATTGCCTGACATCAGAG 241


RESULT 10
AU147973/c
LOCUS       AU147973                462 bp    mRNA    linear   EST 05-AUG-2002
DEFINITION  AU147973 MAMMA1 Homo sapiens cDNA clone MAMMA1002282 3', mRNA
            sequence.
ACCESSION   AU147973
VERSION     AU147973.1  GI:11009494
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T., Sugano
            ,S. and Isogai,T.).
  TITLE     HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S.,
            Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura
            ,Y., Nagai,T., Sugano,S., Isogai,T.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MAMMA1002282"
                     /clone_lib="MAMMA1"
                     /tissue_type="mammary gland"
                     /note="Vector: pME18SFL3"
BASE COUNT      112 a    127 c     99 g    115 t      9 others
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 9;  Length 462;
  Best Local Similarity   94.4%;  Pred. No. 2.4e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAG 18
              |||||||||| ||||||
Db        225 GATTGCCTGAGGTCAGAG 208


RESULT 11
AA524774/c
LOCUS       AA524774                468 bp    mRNA    linear   EST 05-AUG-1997
DEFINITION  nh33f04.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:954175
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA524774
VERSION     AA524774.1  GI:2265702
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 468)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 421   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 384.
FEATURES             Location/Qualifiers
     source          1. .468
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:954175"
                     /clone_lib="NCI_CGAP_Pr3"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected cells histologically-determined to be
                     fully malignant prostate cancer cells.  Double-stranded
                     cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                     applied to the cDNA with an adaptor-specific primer, and
                     the resulting PCR product subcloned into pAMP10 by the
                     UDG-cloning method (Life Technologies).  Average insert
                     size is 600 bp. NOTE: Not directionally cloned.  This
                     library was constructed by  David Krizman."
BASE COUNT      103 a    119 c    101 g    145 t
```

```
ORIGIN

  Query Match            82.0%;  Score 16.4;  DB 9;  Length 468;
  Best Local Similarity  94.4%;  Pred. No. 2.4e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAG 18
              |||||||||| |||||||
Db        240 GATTGCCTGAGGTCAGAG 223


RESULT 12
BF439459/c
LOCUS       BF439459                476 bp    mRNA    linear   EST 30-MAR-2001
DEFINITION  nab64g11.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
            IMAGE:3272660 3', mRNA sequence.
ACCESSION   BF439459
VERSION     BF439459.1  GI:11451976
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 476)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .476
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3272660"
                     /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from five normalized
                     libraries were mixed, and ss circles were made in vitro.
                     Following HAP purification, this DNA was used as tracer in
                     a subtractive hybridization reaction. The driver was
                     PCR-amplified cDNAs from pools of 5,000 clones made from
                     the same 5 libraries. The pools consisted of the following
                     libraries and cloneIDs: Soares NbHSF pool 1:
                     309384-310919, 323208-325895 Soares Nb2HP pool 1:
                     145032-147335, 147720-148103, 148872-149255, 15002 -
                     150407, 151176-152327 Soares Nb2HF8-9W pool 1:
                     758280-760583, 772104-774407 Soares NbHPA pool 1:
                     304776-306311, 320136-322823, 326280-326663 Soares NbHOT
                     pool 1:  723720-726407, 739080-740999 Subtraction by Bento
                     Soares and M. Fatima Bonaldo."
BASE COUNT      151 a     98 c     88 g    139 t
ORIGIN

  Query Match            82.0%;  Score 16.4;  DB 12;  Length 476;
  Best Local Similarity  94.4%;  Pred. No. 2.4e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAG 18
              |||||||||| |||||||
Db        346 GATTGCCTGACATCAGAG 329


RESULT 13
AQ617221/c
LOCUS       AQ617221                570 bp    DNA     linear   GSS 15-JUN-1999
DEFINITION  HS_5152_B1_B01_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=728 Col=1 Row=D, DNA sequence.
ACCESSION   AQ617221
VERSION     AQ617221.1  GI:5078497
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 728  row: D  column: 1
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 570.
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=728 Col=1 Row=D"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      170 a     111 c     113 g     167 t       9 others
ORIGIN

  Query Match            82.0%;  Score 16.4;  DB 17;  Length 570;
  Best Local Similarity  94.4%;  Pred. No. 2.6e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              ||||||||| ||||||||
Db        147 TTGCCTGACCTCAGAGAG 130


RESULT 14
AQ392594
LOCUS       AQ392594                618 bp    DNA     linear   GSS 06-MAR-1999
DEFINITION  CITBI-E1-2546E7.TF CITBI-E1 Homo sapiens genomic clone 2546E7, DNA
            sequence.
ACCESSION   AQ392594
VERSION     AQ392594.1  GI:4363617
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 618)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
```

```
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CITBI-E1-2546E7.TR
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .618
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2546E7"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
BASE COUNT      177 a    148 c    142 g    151 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 17;  Length 618;
  Best Local Similarity   94.4%;  Pred. No. 2.7e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              |||||||| |||||||||
Db        435 TTGCCTGAGGTCAGAGAG 452


RESULT 15
AQ884463/c
LOCUS       AQ884463                630 bp    DNA     linear   GSS 09-NOV-1999
DEFINITION  HS_5510_A2_H02_SP6E RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=9278 Col=4 Row=O, DNA sequence.
ACCESSION   AQ884463
VERSION     AQ884463.1  GI:6315930
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 9278  row: O  column: 4
            Seq primer: SP6
            Class: BAC ends
            High quality sequence stop: 630.
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=9278 Col=4 Row=O"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      189 a    117 c    125 g    181 t     18 others
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 17;  Length 630;
  Best Local Similarity   94.4%;  Pred. No. 2.8e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              |||||||| |||||||||
Db        213 TTGCCTGACCTCAGAGAG 196


Search completed: June 26, 2003, 22:12:14
Job time : 1534.13 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 10:58:25 ; Search time 30.2174 Seconds
(without alignments)
202.980 Million cell updates/sec

Title: US-09-355-254F-8
Perfect score: 20
Sequence: 1 gattgcctgacgtcagagag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	2	US-09-133-774-10	Sequence 10, Appl
	2	20	100.0	20	3	US-09-303-862-10	Sequence 10, Appl
	3	20	100.0	27	1	US-08-210-880B-2	Sequence 2, Appli
	4	20	100.0	27	2	US-08-771-411-2	Sequence 2, Appli
	5	16.8	84.0	12394	4	US-09-488-856A-10	Sequence 10, Appl
	6	16.4	82.0	32	4	US-09-215-098-1	Sequence 1, Appli
	7	16.4	82.0	50	1	US-08-171-389-451	Sequence 451, App
	8	16.4	82.0	50	1	US-08-123-936-451	Sequence 451, App
	9	16.4	82.0	50	2	US-08-475-228A-451	Sequence 451, App
	10	16.4	82.0	50	3	US-08-482-080A-451	Sequence 451, App
	11	16.4	82.0	50	4	US-09-354-947-451	Sequence 451, App
	12	16.4	82.0	50	5	PCT-US93-12388-451	Sequence 451, App
c	13	15.8	79.0	5134	2	US-08-310-912A-157	Sequence 157, App
c	14	15.8	79.0	5134	4	US-09-301-085-157	Sequence 157, App
c	15	15.8	79.0	5134	5	PCT-US95-04589-157	Sequence 157, App
c	16	15.8	79.0	5475	2	US-08-680-327-1	Sequence 1, Appli
c	17	15.8	79.0	5475	4	US-09-228-246-3	Sequence 3, Appli
c	18	15.8	79.0	10968	2	US-08-680-327-2	Sequence 2, Appli
c	19	15.8	79.0	10968	4	US-09-228-246-1	Sequence 1, Appli
	20	15.2	76.0	2559	4	US-09-118-408-43	Sequence 43, Appl
	21	15.2	76.0	2559	4	US-09-506-855-43	Sequence 43, Appl
	22	15	75.0	1974	1	US-08-625-322-1	Sequence 1, Appli
c	23	14.8	74.0	5687	4	US-09-221-017B-368	Sequence 368, App
	24	14.8	74.0	48974	4	US-08-920-422-17	Sequence 17, Appl
	25	14.4	72.0	1243	4	US-09-257-179-30	Sequence 30, Appl
c	26	14.4	72.0	4793	4	US-09-561-497-10	Sequence 10, Appl
	27	14.4	72.0	5375	3	US-08-757-223-7	Sequence 7, Appli
c	28	14.4	72.0	7676	1	US-08-451-777A-7	Sequence 7, Appli
c	29	14.4	72.0	7676	2	US-08-451-778A-7	Sequence 7, Appli
c	30	14.4	72.0	7676	2	US-08-998-208-7	Sequence 7, Appli
c	31	14.4	72.0	7676	5	PCT-US95-06743-7	Sequence 7, Appli
c	32	14.4	72.0	15297	4	US-09-817-180-3	Sequence 3, Appli
	33	14.4	72.0	16063	4	US-09-801-052-3	Sequence 3, Appli
	34	14.4	72.0	99500	4	US-09-798-096-10	Sequence 10, Appl
	35	14.4	72.0	162450	4	US-09-345-882-1	Sequence 1, Appli
c	36	14.4	72.0	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	37	14.4	72.0	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	38	14.2	71.0	41	3	US-08-813-507-78	Sequence 78, Appl
c	39	14.2	71.0	41	4	US-09-464-453-78	Sequence 78, Appl
	40	14.2	71.0	562	4	US-09-449-285A-16	Sequence 16, Appl
c	41	14.2	71.0	1017	4	US-09-330-611-5	Sequence 5, Appli
c	42	14.2	71.0	2100	1	US-08-332-576-1	Sequence 1, Appli
c	43	14.2	71.0	2100	5	PCT-US95-13672-1	Sequence 1, Appli
	44	14.2	71.0	29598	4	US-09-341-587-6	Sequence 6, Appli
	45	13.8	69.0	71	4	US-08-870-930-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-09-133-774-10
; Sequence 10, Application US/09133774B
; Patent No. 5962636
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 5962636el Peptides Capable of Modulating Inflammatory Hear
; TITLE OF INVENTION: Disease
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/133,774B
; CURRENT FILING DATE: 1998-08-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
; FEATURE:
; OTHER INFORMATION: An oligonucleotide derived from the DNA encoding a
; OTHER INFORMATION: 60 kDa cysteine rich outer membrane protein from
; OTHER INFORMATION: Chlamydia trachomatis.
US-09-133-774-10

Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          1 GATTGCCTGACGTCAGAGAG 20
```

RESULT 2
US-09-303-862-10
; Sequence 10, Application US/09303862
; Patent No. 6034230
; GENERAL INFORMATION:
; APPLICANT: Bachmaier, Kurt
; APPLICANT: Hessel, Andrew J.
; APPLICANT: Neu M.D., Nikolaus
; APPLICANT: Penninger, Josef M.
; TITLE OF INVENTION: No. 6034230el Peptides Capable of Modulating Inflammatory Hear
; TITLE OF INVENTION: Disease
; FILE REFERENCE: A-536
; CURRENT APPLICATION NUMBER: US/09/303,862
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: 09/133,774

```
;  EARLIER FILING DATE: 1998-08-12
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Chlamydia trachomatis
;   FEATURE:
;   OTHER INFORMATION: An oligonucleotide derived from the DNA encoding a
;   OTHER INFORMATION: 60 kDa cysteine rich outer membrane protein from
;   OTHER INFORMATION: Chlamydia trachomatis.
US-09-303-862-10

  Query Match           100.0%;  Score 20;  DB 3;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 0.059;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          1 GATTGCCTGACGTCAGAGAG 20


RESULT 3
US-08-210-880B-2
; Sequence 2, Application US/08210880B
; Patent No. 5641486
;  GENERAL INFORMATION:
;    APPLICANT:  HINRICHS, STEVEN H.
;    APPLICANT:  ORTEN, DANA J.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR INHIBITING
;    TITLE OF INVENTION:  CREB/ATF1 FAMILY DETERMINED TRANSCRIPTION
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HENDERSON & STURM
;      STREET:  1125 S. 103RD ST., #330
;      CITY:  OMAHA
;      STATE:  NE
;      COUNTRY:  US
;      ZIP:  68124
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/210,880B
;      FILING DATE:  18-MAR-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JONDLE, ROBERT J.
;      REGISTRATION NUMBER:  33,915
;      REFERENCE/DOCKET NUMBER:  63066
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  402-398-9000
;      TELEFAX:  402-398-9005
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-210-880B-2

  Query Match           100.0%;  Score 20;  DB 1;  Length 27;
  Best Local Similarity  100.0%;  Pred. No. 0.062;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23


RESULT 4
US-08-771-411-2
; Sequence 2, Application US/08771411
; Patent No. 5844096
;  GENERAL INFORMATION:
;    APPLICANT:  HINRICHS, STEVEN H.
;    APPLICANT:  ORTEN, DANA J.
;    TITLE OF INVENTION:  METHOD AND COMPOSITION FOR INHIBITING
;    TITLE OF INVENTION:  CREB/ATF1 FAMILY DETERMINED TRANSCRIPTION
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HENDERSON & STURM
;      STREET:  1125 S. 103RD ST., #330
;      CITY:  OMAHA
;      STATE:  NE
;      COUNTRY:  US
;      ZIP:  68124
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/771,411
;      FILING DATE:  20-DEC-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/210,880
;      FILING DATE:  18-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JONDLE, ROBERT J.
;      REGISTRATION NUMBER:  33,915
;      REFERENCE/DOCKET NUMBER:  63066
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  402-398-9000
;      TELEFAX:  402-398-9005
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-771-411-2

  Query Match           100.0%;  Score 20;  DB 2;  Length 27;
  Best Local Similarity  100.0%;  Pred. No. 0.062;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23


RESULT 5
US-09-488-856A-10
; Sequence 10, Application US/09488856A
; Patent No. 6316259
; GENERAL INFORMATION:
;  APPLICANT:      Brett P. Monia
;  APPLICANT:  Robert McKay
;  APPLICANT:  Madeline M. Butler
;  APPLICANT:  Jacqueline Wyatt
;  TITLE OF INVENTION:      ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA
;  FILE REFERENCE:      RTS-0115
;  CURRENT APPLICATION NUMBER: US/09/488,856A
;  CURRENT FILING DATE:  2000-01-21
;  NUMBER OF SEQ ID NOS:      88
; SEQ ID NO 10
;   LENGTH: 12394
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (115)...(397)
;   NAME/KEY: CDS
;   LOCATION: (2438)...(2625)
;   NAME/KEY: CDS
;   LOCATION: (5639)...(5722)
;   NAME/KEY: CDS
;   LOCATION: (5864)...(5974)
;   NAME/KEY: CDS
;   LOCATION: (7902)...(8032)
;   NAME/KEY: CDS
;   LOCATION: (8121)...(8227)
;   NAME/KEY: CDS
;   LOCATION: (9197)...(9294)
;   NAME/KEY: CDS
;   LOCATION: (9375)...(9470)
;   NAME/KEY: CDS
;   LOCATION: (9898)...(10084)
;   NAME/KEY: CDS
;   LOCATION: (10431)...(10523)
;   NAME/KEY: CDS
;   LOCATION: (11713)...(11786)
US-09-488-856A-10

  Query Match             84.0%;  Score 16.8;  DB 4;  Length 12394;
  Best Local Similarity   90.0%;  Pred. No. 7;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              |||||||||| |||| ||||
Db       2949 GATTGCCTGAGGTCATAGAG 2968


RESULT 6
US-09-215-098-1
; Sequence 1, Application US/09215098
; Patent No. 6194632
; GENERAL INFORMATION:
;  APPLICANT:  Leiden, Jeffery M
;  TITLE OF INVENTION:  DILATED CARDIOMYOPATHY IN TRANSGENIC MICE EXPRESSING A
;  TITLE OF INVENTION:  DOMINANT-NEGATIVE CREB TRANSCRIPTION FACTOR IN THE
;  TITLE OF INVENTION:  HEART
;  FILE REFERENCE:  9189-4
;  CURRENT APPLICATION NUMBER: US/09/215,098
;  CURRENT FILING DATE:  1998-12-18
;  PRIOR APPLICATION NUMBER:  60/068,011
;  PRIOR FILING DATE:  1997-12-18
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Oligonucleotide
;   OTHER INFORMATION: containing the CREB site from the somatostatin
;   OTHER INFORMATION: promoter
US-09-215-098-1

  Query Match             82.0%;  Score 16.4;  DB 4;  Length 32;
  Best Local Similarity   94.4%;  Pred. No. 5.1;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              ||| ||||||||||||||
Db         11 TTGGCTGACGTCAGAGAG 28


RESULT 7
US-08-171-389-451
; Sequence 451, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  451:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human somatostatin I gene
US-08-171-389-451

  Query Match             82.0%;  Score 16.4;  DB 1;  Length 50;
  Best Local Similarity   94.4%;  Pred. No. 5.4;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              | ||||||||||||||||
Db          1 TAGCCTGACGTCAGAGAG 18


RESULT 8
US-08-123-936-451
; Sequence 451, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
```

```
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  451:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human somatostatin I gene
US-08-123-936-451

  Query Match             82.0%;  Score 16.4;  DB 1;  Length 50;
  Best Local Similarity   94.4%;  Pred. No. 5.4;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              | ||||||||||||||||
Db          1 TAGCCTGACGTCAGAGAG 18


RESULT 9
US-08-475-228A-451
; Sequence 451, Application US/08475228A
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,228A
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A.
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0175.21/G19P3D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  451:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human somatostatin I gene
US-08-475-228A-451

  Query Match             82.0%;  Score 16.4;  DB 2;  Length 50;
  Best Local Similarity   94.4%;  Pred. No. 5.4;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              | ||||||||||||||||
Db          1 TAGCCTGACGTCAGAGAG 18


RESULT 10
US-08-482-080A-451
; Sequence 451, Application US/08482080A
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  451:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human somatostatin I gene
US-08-482-080A-451

  Query Match             82.0%;  Score 16.4;  DB 3;  Length 50;
  Best Local Similarity   94.4%;  Pred. No. 5.4;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              | ||||||||||||||||
Db          1 TAGCCTGACGTCAGAGAG 18


RESULT 11
US-09-354-947-451
; Sequence 451, Application US/09354947
; Patent No. 6384208
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/354,947
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/482,080
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  451:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human somatostatin I gene
US-09-354-947-451

  Query Match             82.0%;  Score 16.4;  DB 4;  Length 50;
  Best Local Similarity   94.4%;  Pred. No. 5.4;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              | ||||||||||||||||
Db          1 TAGCCTGACGTCAGAGAG 18


RESULT 12
PCT-US93-12388-451
; Sequence 451, Application PC/TUS9312388
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  451:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human somatostatin I gene
PCT-US93-12388-451

  Query Match             82.0%;  Score 16.4;  DB 5;  Length 50;
  Best Local Similarity   94.4%;  Pred. No. 5.4;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 TTGCCTGACGTCAGAGAG 20
              | ||||||||||||||||
Db          1 TAGCCTGACGTCAGAGAG 18


RESULT 13
US-08-310-912A-157/c
; Sequence 157, Application US/08310912A
; Patent No. 5981730
;  GENERAL INFORMATION:
;    APPLICANT:  Ausubel, Frederick M.
;    APPLICANT:  Staskawicz, Brian J.
;    APPLICANT:  Brent, Andrew F.
;    APPLICANT:  Dahlbeck, Douglas
;    APPLICANT:  Katagiri, Fumiaki
;    APPLICANT:  Kunkel, Barbara N.
;    APPLICANT:  Mindrinos, Michael N.
;    APPLICANT:  Yu, Guo-Liang
;    TITLE OF INVENTION:  RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  208
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2904
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/310,912A
;      FILING DATE:  September 22, 1994
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/227,360
;      FILING DATE:  April 13, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lech, Karen F.
;      REGISTRATION NUMBER:  35,238
;      REFERENCE/DOCKET NUMBER:  00786/254001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  100254
;  INFORMATION FOR SEQ ID NO:  157:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5134 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-310-912A-157

  Query Match             79.0%;  Score 15.8;  DB 2;  Length 5134;
  Best Local Similarity   89.5%;  Pred. No. 21;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 ATTGCCTGACGTCAGAGAG 20
              ||||| |||| ||||||||
Db        721 ATTGCTTGACATCAGAGAG 703


RESULT 14
US-09-301-085-157/c
; Sequence 157, Application US/09301085
; Patent No. 6262248
; GENERAL INFORMATION:
;  APPLICANT: Ausubel, Frederick M.
;  APPLICANT: Staskawicz, Brian J.
;  APPLICANT: Brent, Andrew F.
;  APPLICANT: Dahlbeck, Douglas
;  APPLICANT: Katagiri, Fumiaki
;  APPLICANT: Kunkel, Barbara N.
;  APPLICANT: Mindrinos, Michael N.
;  APPLICANT: Yu, Guo-Liang
;  TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
;  TITLE OF INVENTION: DETECTION  METHODS
;  FILE REFERENCE: 00786/254002
;  CURRENT APPLICATION NUMBER: US/09/301,085
;  CURRENT FILING DATE: 1999-04-28
;  EARLIER APPLICATION NUMBER: 08/310,912
;  EARLIER FILING DATE: 1994-09-22
;  EARLIER APPLICATION NUMBER: 08/227,360
;  EARLIER FILING DATE: 1994-04-13
;  NUMBER OF SEQ ID NOS: 208
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 157
;   LENGTH: 5134
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-301-085-157

  Query Match             79.0%;  Score 15.8;  DB 4;  Length 5134;
  Best Local Similarity   89.5%;  Pred. No. 21;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 ATTGCCTGACGTCAGAGAG 20
              ||||| |||| ||||||||
Db        721 ATTGCTTGACATCAGAGAG 703


RESULT 15
PCT-US95-04589-157/c
; Sequence 157, Application PC/TUS9504589
;  GENERAL INFORMATION:
;    APPLICANT:  Ausubel, Frederick M.
;    APPLICANT:  Staskawicz, Brian J.
;    APPLICANT:  Brent, Andrew F.
;    APPLICANT:  Dahlbeck, Douglas
;    APPLICANT:  Katagiri, Fumiaki
;    APPLICANT:  Kunkel, Barbara N.
;    APPLICANT:  Mindrinos, Michael N.
;    APPLICANT:  Yu, Guo-Liang
;    TITLE OF INVENTION:  RPS2 GENE AND USES THEREOF
;    NUMBER OF SEQUENCES:  201
;    CORRESPONDENCE ADDRESS:
```

```
;       ADDRESSEE:  Fish & Richardson
;       STREET:  225 Franklin Street Suite 3100
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02110-2904
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/04589
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/227,360
;       FILING DATE:  13-APR-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00786/230001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  100254
;   INFORMATION FOR SEQ ID NO:  157:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5134 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
PCT-US95-04589-157

  Query Match             79.0%;  Score 15.8;  DB 5;  Length 5134;
  Best Local Similarity   89.5%;  Pred. No. 21;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 ATTGCCTGACGTCAGAGAG 20
              ||||| |||| ||||||||
Db        721 ATTGCTTGACATCAGAGAG 703


Search completed: June 26, 2003, 16:20:18
Job time : 37.2888 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 09:43:36 ; Search time 226.087 Seconds
(without alignments)
199.215 Million cell updates/sec

Title: US-09-355-254F-8
Perfect score: 20
Sequence: 1 gattgcctgacgtcagagag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	19	AAV46000	Immune adjuvant AP
	2	20	100.0	20	20	AAZ28189	Chlamydia trachoma
	3	20	100.0	20	21	AAZ99172	Inflammatory cardi
	4	20	100.0	20	24	AAL39177	Murine Toll-like r
	5	20	100.0	26	20	AAX76047	cAMP response elem
	6	20	100.0	27	18	AAT85832	CRE oligonucloide
	7	20	100.0	27	20	AAV82454	ATF comp oligonucl
	8	20	100.0	27	20	AAV08336	CRE element coding
	9	20	100.0	27	22	AAI70581	Transcription fact
	10	20	100.0	27	22	AAH77396	Cyclic AMP respons
c	11	20	100.0	27	22	AAH77397	Cyclic AMP respons
	12	20	100.0	27	22	AAF87956	Cyclic-AMP respons
c	13	20	100.0	27	22	AAF87957	Cyclic-AMP respons
	14	20	100.0	27	22	AAF76267	cAMP response elem
	15	20	100.0	27	24	ABA92274	CRE binding site o
	16	20	100.0	27	24	ABA05538	Cyclic-AMP respons
	17	19	95.0	20	19	AAV45997	Immune adjuvant CR
	18	16.8	84.0	12394	22	AAD14749	Human glycogen syn
	19	16.4	82.0	28	24	ABK14052	Cyclic AMP respons
	20	16.4	82.0	32	22	AAF77813	CREB probe derived
	21	16.4	82.0	37	19	AAV04084	Somatostatin gene
	22	16.4	82.0	50	15	AAQ69701	Human somatostatin
	23	16.4	82.0	50	18	AAT64163	Human somatostatin
	24	16.4	82.0	50	20	AAX17451	Test sequence from
	25	16.4	82.0	50	24	ABK82942	DNA binding molecu
c	26	16.4	82.0	462	22	AAH10445	Human cDNA clone (
	27	16.4	82.0	723	20	AAZ41392	Human normal pancr
	28	16.4	82.0	2443	22	AAH17546	Human cDNA sequenc
	29	16.4	82.0	2564	24	ABI99414	Mouse ischaemic co
	30	16	80.0	462	23	ABV49687	Human prostate exp
	31	15.8	79.0	20	19	AAV45999	Immune adjuvant CR
	32	15.8	79.0	917	22	AAH94473	Human foetal cDNA,
c	33	15.8	79.0	2662	22	AAK86924	Human immune/haema
c	34	15.8	79.0	2662	22	AAK86925	Human immune/haema
c	35	15.8	79.0	2669	22	AAK86923	Human immune/haema
c	36	15.8	79.0	3648	24	AAD37410	Human phospholipas
c	37	15.8	79.0	5475	19	AAV17777	Tomato Prf cDNA.
	38	15.8	79.0	5592	22	AAF30935	Spodoptera frugipe
c	39	15.8	79.0	10968	19	AAV17789	Tomato Prf genomic
	40	15.4	77.0	123	22	ABA76163	Human foetal liver
	41	15.4	77.0	123	22	ABA40714	Probe #19180 for g
	42	15.4	77.0	123	22	AAK24826	Human brain expres
	43	15.4	77.0	123	22	AAI27849	Probe #17782 for g
	44	15.4	77.0	123	22	AAI56827	Probe #25513 used
	45	15.4	77.0	123	24	ABS24317	Human genome-deriv

ALIGNMENTS

RESULT 1
AAV46000
ID AAV46000 standard; DNA; 20 BP.
XX
AC AAV46000;
XX
DT 16-OCT-1998 (first entry)
XX
DE Immune adjuvant AP-1 #1.
XX
KW Immune system; adjuvant; vaccine; cancer; prophylactic; pathogenicity;
KW modulator; tolerance; regulator; helper cell; antigen; immunoglobulin;
KW Ig class; autoimmune response; T-cell; B-cell; tumour; ss.
XX
OS Class Bacteria.
XX
PN EP855184-A1.
XX
PD 29-JUL-1998.
XX
PF 23-JAN-1997; 97EP-0101019.
XX
PR 23-JAN-1997; 97EP-0101019.
XX
PA (HEEG/) HEEG K.
PA (LIPF/) LIPFORD G B.
PA (WAGN/) WAGNER H.
XX
PI Heeg K, Lipford GB, Wagner H;
XX
DR WPI; 1998-389630/34.
XX

```
PT   Antigenic composition comprises polynucleotide fragment and antigen
PT   - used as vaccine to treat or prevent e.g. cancer or pathogen
PT   infections and to modulate immune response e.g. tolerance break and
PT   regulation of TH1/TH2 cells
XX
PS   Example 3; Page 7; 28pp; English.
XX
CC   AAV45993-V46019 are fragments of bacterial polynucleotides which are
CC   used as immune adjuvants for inclusion into vaccines to treat cancer and
CC   for prophylaxis and/or treatment of conditions caused by pathogenic
CC   micro-organisms. The polynucleotide is used for modulation of an immune
CC   response and the modulation is selected from the group break of
CC   tolerance, regulation of TH1/TH2 helper cell responses, switch of Ig
CC   classes, treatment of autoimmune responses and induction of tolerances.
CC   DNA oligomers are used to enhance the reactivity of immune cells to
CC   viral, bacterial and parasitic antigens, to break tolerance in anergic T
CC   and B cells e.g. against tumour antigens, as adjuvants in vaccination
CC   against tumour-defined antigens and immunostimulatory substances in an
CC   immune response against tumours and to suppress immune reactions of the
CC   innate and acquired immune system. The composition is inexpensive and
CC   stable and does not cause lethal shock, which happens with prior art
CC   bacterial sequences.
XX
SQ   Sequence 20 BP; 5 A; 4 C; 7 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 19;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.48;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          1 GATTGCCTGACGTCAGAGAG 20


RESULT 2
AAZ28189
ID   AAZ28189 standard; DNA; 20 BP.
XX
AC   AAZ28189;
XX
DT   20-DEC-1999  (first entry)
XX
DE   Chlamydia trachomatis outer membrane protein gene-derived CpG oligo 2.
XX
KW   Heart disease; inflammatory; autoimmune; cardiomyopathy; adjuvant;
KW   CpG motif; vaccine; ds.
XX
OS   Synthetic.
OS   Chlamydia trachomatis.
XX
FH   Key             Location/Qualifiers
FT   modified_base    1..20
FT                    /*tag= a
FT                    /mod_base= OTHER
FT                    /note= "OTHER = phosphorothioate linkage"
XX
PN   US5962636-A.
XX
PD   05-OCT-1999.
XX
PF   12-AUG-1998;   98US-0133774.
XX
PR   12-AUG-1998;   98US-0133774.
XX
PA   (AMGE-) AMGEN CANADA INC.
XX
PI   Bachmaier K,  Hessel AJ,  Penninger JM,  Neu N;
XX
DR   WPI; 1999-589735/50.
XX
PT   Peptides that induce or suppress inflammatory cardiomyopathy  -
XX
PS   Example 2; Column 25; 17pp; English.
XX
CC   This sequence represents DNA encoding Chlamydia trachomatis 60 kD outer
CC   membrane protein (OMP) gene-derived CpG oligonucleotide 2. This
CC   oligonucleotide contains a CpG motif. It was tested for its ability to
CC   act as an adjuvant for the M7A-alpha peptide (AAY42723), which can
CC   induce inflammatory cardiomyopathy (ICM) in mice. It was found to act as
CC   a potent immunostimulator, whereas a oligonucleotide from the same
CC   source which did not contain a CpG motif (AAZ28193) was hardly effective
CC   as an adjuvant. Inflammatory cardiomyopathy peptides (AAY42723,
CC   AAY42725-Y42731) can be used with such an adjuvant and an excipient in a
CC   vaccine for decreasing ICM.
XX
SQ   Sequence 20 BP; 5 A; 4 C; 7 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 20;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.48;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          1 GATTGCCTGACGTCAGAGAG 20


RESULT 3
AAZ99172
ID   AAZ99172 standard; DNA; 20 BP.
XX
AC   AAZ99172;
XX
DT   21-JUN-2000  (first entry)
XX
DE   Inflammatory cardiomyopathy immunostimulatory oligonucleotide #1.
XX
KW   Cardiant; murine alpha myosin heavy chain; inflammatory myocarditis;
KW   autoimmune inflammatory cardiomyopathy; Chlamydia; antibody; vaccine;
KW   hybridization probe; immunostimulatory; ss.
XX
OS   Synthetic.
XX
PN   US6034230-A.
XX
PD   07-MAR-2000.
XX
PF   03-MAY-1999;   99US-0303862.
XX
PR   12-AUG-1998;   98US-0133774.
XX
PA   (AMGE-) AMGEN CANADA INC.
XX
PI   Neu N,  Penninger JM,  Bachmaier K,  Hessel AJ;
XX
DR   WPI; 2000-255712/22.
XX
PT   DNA molecules encoding novel myocardial peptides used for inhibiting
PT   and inducing inflammatory cardiomyopathy in vivo  -
XX
PS   Disclosure; Column 17; 17pp; English.
XX
CC   The invention relates to the isolation of sequences coding for peptide
CC   sequences derived from bacteria and viruses which may cause inflammatory
CC   cardiomyopathy. The peptide sequences are searched based on the sequence
CC   of the M7A peptides derived from the murine alpha myosin heavy chain
CC   polypeptide. The consensus sequence of the murine M7A-alpha/beta peptides
CC   (Y83813) was used to search the PIR public database for similar bacterial
CC   and viral sequences able to cause inflammatory cardiomyopathy. The screen
CC   isolated the peptides Y83814-Y83819 and their corresponding coding
CC   sequences Z99164-Z99169. The peptides encoded by the DNAs are used, alone
CC   or in conjunction with other therapeutics, for inducing or inhibiting
CC   inflammatory cardiomyopathy in vivo, where the cardiomyopathy is
CC   autoimmune inflammatory cardiomyopathy, and inflammatory cardiomyopathy
CC   caused by Chlamydia or other bacterial or viral infections that cause
CC   inflammatory cardiomyopathy. The oligonucleotides Z99172-Z99176 were
CC   shown to increase the immunogenicity of the immunostimulatory peptides
```

```
CC   when injected simultaneously. The peptides may also be used for
CC   increasing inflammatory myocarditis in a mammal. Antibodies against the
CC   peptides and the peptides themselves are used for measuring the risk of
CC   inflammatory cardiomyopathy in a mammal. The peptides may also be used
CC   in vaccines. Nucleic acids encoding the peptides may be used as
CC   hybridization probes, e.g. in diagnostic assays to test for the
CC   presence of Chlamydia DNA.
XX
SQ   Sequence 20 BP; 5 A; 4 C; 7 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.48;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          1 GATTGCCTGACGTCAGAGAG 20


RESULT 4
AAL39177
ID   AAL39177 standard; DNA; 20 BP.
XX
AC   AAL39177;
XX
DT   05-SEP-2002  (first entry)
XX
DE   Murine Toll-like receptor related CpG DNA SEQ ID No 52.
XX
KW   Murine Toll-like receptor; TLR9; TLR7; TLR8; ISNA; ds.
XX
OS   Unidentified.
XX
PN   WO200222809-A2.
XX
PD   21-MAR-2002.
XX
PF   17-SEP-2001; 2001WO-US29229.
XX
PR   15-SEP-2000; 2000US-233035P.
PR   23-JAN-2001; 2001US-263657P.
PR   17-MAY-2001; 2001US-291726P.
PR   22-JUN-2001; 2001US-300210P.
XX
PA   (COLE-) COLEY PHARM GMBH.
XX
PI   Bauer S,  Lipford G,  Wagner H;
XX
DR   WPI; 2002-393964/42.
XX
PT   New isolated murine Toll-like receptor (TLR)9, TLR7, TLR8 polypeptides,
PT   useful for identifying species specificity of immunostimulatory nucleic
PT   acid and identifying immunostimulatory nucleic acids
XX
PS   Disclosure; Page 76; 195pp; English.
XX
CC   The invention relates to isolated murine Toll-like receptors (TLR)9,
CC   TLR7 and TLR8 polypeptides. These polypeptides comprise fully defined
CC   sequences of 1032, 1050 or 1032 amino acids as given in specification, or
CC   their fragments, where TLR9, TLR7 and TLR8 polypeptides or their
CC   fragments have an amino acid sequence which is identical to human TLR9,
CC   TLR7 or TLR8 polypeptides or their fragment except for at least one amino
CC   acid of a murine TLR polypeptide. The isolated nucleic acids of the
CC   invention are useful for inhibiting TLR9 signalling activity in a cell.
CC   TLR7, TLR8 and TLR9 polypeptides are useful for identifying nucleic acid
CC   molecules which interact with a TLR polypeptide or its fragment. The
CC   TLR7, TLR8 or TLR9 polypeptides are also useful for identifying ISNA. The
CC   TLR7, TLR8 and TLR9 polypeptides are also useful for comparing TLR9
CC   signalling activity of a test compound (that is not a nucleic acid, and
CC   is a polypeptide or a part of a combinatorial library of compounds) with
CC   an ISNA. The TLR7, TLR8 and TLR9 polypeptides are also useful for
CC   identifying species specificity of an ISNA. The isolated nucleic acids of
CC   the invention are useful as probes or primers. This polynucleotide
CC   sequence represents DNA relating to the isolated Toll-like receptors of
CC   the invention.
XX
SQ   Sequence 20 BP; 5 A; 4 C; 7 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 24;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.48;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          1 GATTGCCTGACGTCAGAGAG 20


RESULT 5
AAX76047
ID   AAX76047 standard; DNA; 26 BP.
XX
AC   AAX76047;
XX
DT   30-JUL-1999  (first entry)
XX
DE   cAMP response element oligonucleotide SEQ ID NO:15.
XX
KW   CRE; cAMP response element; transcription factor decoy; cis-element;
KW   tumour growth inhibitor; palindrome; hairpin; cancer; metabolism;
KW   gene transcription regulation; inhibiting proliferation; ds.
XX
OS   Synthetic.
XX
PN   WO9926634-A1.
XX
PD   03-JUN-1999.
XX
PF   23-NOV-1998;   98WO-US25307.
XX
PR   24-NOV-1997;   97US-0977643.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Cho-Chung YS;
XX
DR   WPI; 1999-347612/29.
XX
PT   Nucleic acids that compete with response elements for transcription
PT   factors
XX
PS   Example 10; Page 54; 83pp; English.
XX
CC   The present invention describes a composition (A) comprising one or more
CC   nucleic acids (I) that compete with cAMP (cyclic adenosine monophosphate)
CC   response element (CRE) enhancer DNA for binding to transcription factors
CC   (TF). (I) are used to regulate gene transcription in cells, in vitro or
CC   in vivo, specifically for inhibiting proliferation of cancer cells, but
CC   possibly also for regulation of metabolism in hepatitis B and other
CC   viruses. HCT-15 human multidrug resistant colon carcinoma cells (2
CC   million) were inoculated subcutaneously into the flank of nude mice,
CC   then the CRE oligonucleotide  5'-TGACGTTCATGACGTTCATGACGTTCA-3' injected
CC   intraperitoneally at doses of 0.1 mg, 5 times per week, once the tumour
CC   had reached 30-50 mg. This treatment resulted in over 85% reduction in
CC   tumour growth, relative to an untreated control. (I) have high affinity
CC   for TF and can inhibit growth of cancer cells without adverse effects on
CC   normal cells (contrast use of antisense RNA). The method does not
CC   require knowledge of the target gene sequence, only of the response
CC   element sequence. The present sequence is used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 26 BP; 8 A; 4 C; 9 G; 5 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 20;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23


RESULT 6
AAT85832
ID   AAT85832 standard; DNA; 27 BP.
XX
AC   AAT85832;
XX
DT   21-NOV-1997  (first entry)
XX
DE   CRE oligonucloide used in gel shift assay.
XX
KW   Activating transcription factor 1; ATF1; CREB; recognition sequence;
KW   cyclic AMP responsive element binding protein; inhibition; binding;
KW   proliferation; virus; cancer; HTLV1; leukaemia; antibody; ss.
XX
OS   Synthetic.
XX
PN   US5641486-A.
XX
PD   24-JUN-1997.
XX
PF   18-MAR-1994;   94US-0210880.
XX
PR   18-MAR-1994;   94US-0210880.
XX
PA   (UYNE-) UNIV NEBRASKA.
XX
PI   Hinrichs SH,  Orten DJ;
XX
DR   WPI; 1997-340900/31.
XX
PT   Inhibiting replication of cancer cells or viruses - with inhibitor
PT   that binds to peptide sequence of activating transcription factor 1
XX
PS   Example 2; Column 6; 17pp; English.
XX
CC   This oligonucleotide sequence corresponds to the cyclic AMP binding
CC   element (CRE) to which members of the activating transcription factor 1
CC   (ATF1)-cyclic AMP responsive element binding protein (CREB) family
CC   of protein bind.  The sequence was used in a gel shift mobility assay to
CC   identify agents which inhibit the binding of ATF1 to its recognition
CC   sequence.  The agents are preferably antibodies, small molecules or
CC   polypeptides, especially the complementarity determining region of
CC   monoclonal antibody MAb4.  The agents cause inhibition of transcription
CC   by dissociating ATF1 from its target gene and thus will prevent
CC   proliferation of e.g. a virus or cancer cell, such as HTLV1-mediated
CC   leukaemic cell proliferation.
XX
SQ   Sequence 27 BP; 8 A; 5 C; 9 G; 5 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 18;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23


RESULT 7
AAV82454
ID   AAV82454 standard; DNA; 27 BP.
XX
AC   AAV82454;
XX
DT   12-APR-1999  (first entry)
XX
DE   ATF comp oligonucleotide used in competition analysis.
XX
KW   Vascular endothelial growth factor; VEGF; human; hypoxia;
KW   vascular disease; tumour; cancer; angiogenesis; wound healing;
KW   therapy; diagnosis; ds.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9856936-A1.
XX
PD   17-DEC-1998.
XX
PF   10-JUN-1998;   98WO-EP03517.
XX
PR   10-JUN-1997;   97EP-0109418.
XX
PA   (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
PI   Damert A,  Plate K,  Risau W;
XX
DR   WPI; 1999-080911/07.
XX
PT   New recombinant DNA - contains sequence that regulates
PT   hypoxia-induced expression, used for, e.g. treatment and diagnosis
PT   of vascular disease
XX
PS   Example 6; Page 41; 80pp; English.
XX
CC   Oligonucleotides hVEGF, hVEGF 5' DEL, AP1M1 and AP1M2 (see
CC   AAV82449-52), and competitor oligonucleotides AP1 comp, ATF comp
CC   and VL30 (see AAV82453-55) were used in electrophoretic mobility
CC   shift assays to determing which transcription factor(s) bind to
CC   the cis-acting element that is involved in the potentiation of
CC   hypoxia inducible factor 1 (HIF-1) mediated hypoxic induction
CC   of vascular endothelial growth factor (VEGF) gene regulatory
CC   sequences.  Experiments were performed using normoxic or hypoxic
CC   C6 cell nuclear extracts.  An AP1 consensus binding site was shown
CC   to compete for DNA-protein-complex formation at potentiating
CC   sequences.  The invention relates to recombinant DNA molecules
CC   comprising regulatory sequences of the VEGF gene, especially the
CC   3' untranslated region (see AAV82439) and promoter (see AAV82440),
CC   being capable of modulating hypoxia inducible expression of a
CC   heterologous DNA in vivo.  Such recombinant DNA molecules, vectors,
CC   host cells and transgenic animals can be used to identify and
CC   develop compounds and methods for diagnosing, treating, preventing
CC   and/or delaying a vascular or tumour disease.
XX
SQ   Sequence 27 BP; 8 A; 5 C; 9 G; 5 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 20;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23


RESULT 8
AAV08336
ID   AAV08336 standard; DNA; 27 BP.
XX
AC   AAV08336;
XX
DT   04-FEB-1999  (first entry)
XX
DE   CRE element coding sequence.
XX
KW   ATF1; activating transcription factor 1; inhibitor; gene transcription;
KW   cell proliferation; cancer cell; human; ds.
XX
OS   Synthetic.
XX
PN   US5844096-A.
```

XX
PD 01-DEC-1998.
XX
PF 20-DEC-1996; 96US-0771411.
XX
PR 18-MAR-1994; 94US-0210880.
PR 20-DEC-1996; 96US-0771411.
XX
PA (UYNE-) UNIV NEBRASKA.
XX
PI Hinrichs SH, Orten DJ;
XX
DR WPI; 1999-044667/04.
XX
PT Inhibitor of activating transcription factor 1 mediated gene
PT transcription - useful as anticancer or antiviral agent
XX
PS Example 2; Column 6; 17pp; English.
XX
CC This sequence represents a CRE element coding sequence. This sequence
CC was used to test the effect of the inhibitory compound of the
CC invention. The inhibitory compound binds to ATF1 residues 167-181 with
CC sufficient affinity to dissociate AFT1 from a gene to which it is bound
CC and thereby prevent transcription of the gene. The inhibitory compound
CC and its derivatives are useful for inhibiting the ATF1-mediated
CC proliferation of cancer cells and viruses, e.g. HTLV I.
XX
SQ Sequence 27 BP; 8 A; 5 C; 9 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 20; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATTGCCTGACGTCAGAGAG 20
||||||||||||||||||||
Db 4 GATTGCCTGACGTCAGAGAG 23

RESULT 9
AAI70581
ID AAI70581 standard; DNA; 27 BP.
XX
AC AAI70581;
XX
DT 21-JAN-2002 (first entry)
XX
DE Transcription factor CREB consensus oligonucleotide.
XX
KW Transcription factor; CREB; screening; detection; quantification;
KW probe; ds.
XX
OS Synthetic.
XX
PN EP1136567-A1.
XX
PD 26-SEP-2001.
XX
PF 24-MAR-2000; 2000EP-0870057.
XX
PR 24-MAR-2000; 2000EP-0870057.
XX
PA (ADAR-) ADVANCED ARRAY TECHNOLOGIES SA.
XX
PI Remacle J, Renard P, Art M;
XX
DR WPI; 2001-640391/74.
XX
PT Screening, detecting or quantifying transcriptional factors in a
PT biological sample comprises contacting the transcriptional factor with
PT a double-stranded DNA sequence bound to an insoluble solid support -
XX
PS Example 4; Page 8; 20pp; English.
XX
CC The present sequence is that of a CREB transcription factor
CC consensus oligonucleotide. Double-stranded probe nucleotide
CC sequences were constructed from 100 bp of a CMV 5' sequence (see
CC AAI70578) linked to this oligonucleotide and used in microwell
CC colorimetric CREB and phospho-CREB assays. The double-stranded
CC probe was biotinylated at its CMV 5' extremity and linked to
CC streptavidin-coated 96-wells plates. The plates were contacted
CC with a nuclear extract of L929 murine fibrosarcoma cells,
CC incubated with anti-CREB or anti-phospho-CREB antibody and then
CC with peroxidase-conjugated antibody. The presence of CREB or
CC phospho-CREB was detected through the action of peroxidase on
CC tetramethylbenzidine. This is an example of the method of the
CC invention, involving the detection of a transcription factor using
CC a double-stranded DNA probe bound to an insoluble solid support at
CC a concentration of at least 0.01 pmole/sq cm of support surface and
CC at a distance of at least 6.8 nm from the surface of the support.
CC The method allows the screening, detection and/or quantification of
CC one or more transcriptional factors, of molecules binding such
CC factors, and of molecules that inhibit such binding, using
CC non-radioactive detection methods.
XX
SQ Sequence 27 BP; 8 A; 5 C; 9 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATTGCCTGACGTCAGAGAG 20
||||||||||||||||||||
Db 4 GATTGCCTGACGTCAGAGAG 23

RESULT 10
AAH77396
ID AAH77396 standard; DNA; 27 BP.
XX
AC AAH77396;
XX
DT 05-NOV-2001 (first entry)
XX
DE Cyclic AMP response element CRE consensus oligonucleotide probe #1.
XX
KW Cancer; human; dithiocarbamate derivative; anticancer agent; probe; ss.
XX
OS Unidentified.
XX
PN US2001016600-A1.
XX
PD 23-AUG-2001.
XX
PF 12-DEC-2000; 2000US-0735205.
XX
PR 08-SEP-1998; 98US-0099390.
PR 08-SEP-1999; 99US-0392122.
PR 05-OCT-2000; 2000US-0679932.
XX
PA (KENN/) KENNEDY T P.
XX
PI Kennedy TP;
XX
DR WPI; 2001-557127/62.
XX
PT Treating cancer, asthma and cancer and reducing hypoxic or ischemic
PT damage comprises administering dithiocarbamate thiolate anion or
PT dithiocarbamate thiolate metal complex -
XX
PS Disclosure; Page 10; 36pp; English.
XX
CC The present invention describes a method of treating cancer, asthma and
CC arthritis and reducing hypoxic or ischaemic damage, involving
CC administering a dithiocarbamate thiolate anion or metal ion complex to
CC the patient. The present sequence is a probe for the cyclic AMP response
CC element CRE, which was described in the exemplification of the invention.

```
XX
SQ   Sequence 27 BP; 8 A; 5 C; 9 G; 5 T; 0 other;

  Query Match            100.0%;  Score 20;  DB 22;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23


RESULT 11
AAH77397/c
ID   AAH77397 standard; DNA; 27 BP.
XX
AC   AAH77397;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Cyclic AMP response element CRE consensus oligonucleotide probe #2.
XX
KW   Cancer; human; dithiocarbamate derivative; anticancer agent; probe; ss.
XX
OS   Unidentified.
XX
PN   US2001016600-A1.
XX
PD   23-AUG-2001.
XX
PF   12-DEC-2000; 2000US-0735205.
XX
PR   08-SEP-1998;   98US-0099390.
PR   08-SEP-1999;   99US-0392122.
PR   05-OCT-2000; 2000US-0679932.
XX
PA   (KENN/) KENNEDY T P.
XX
PI   Kennedy TP;
XX
DR   WPI; 2001-557127/62.
XX
PT   Treating cancer, asthma and cancer and reducing hypoxic or ischemic
PT   damage comprises administering dithiocarbamate thiolate anion or
PT   dithiocarbamate thiolate metal complex -
XX
PS   Disclosure; Page 10; 36pp; English.
XX
CC   The present invention describes a method of treating cancer, asthma and
CC   arthritis and reducing hypoxic or ischaemic damage, involving
CC   administering a dithiocarbamate thiolate anion or metal ion complex to
CC   the patient. The present sequence is a probe for the cyclic AMP response
CC   element CRE, which was described in the exemplification of the invention.
XX
SQ   Sequence 27 BP; 5 A; 9 C; 5 G; 8 T; 0 other;

  Query Match            100.0%;  Score 20;  DB 22;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db         24 GATTGCCTGACGTCAGAGAG 5


RESULT 12
AAF87956
ID   AAF87956 standard; DNA; 27 BP.
XX
AC   AAF87956;
XX
DT   20-JUL-2001  (first entry)
XX
DE   Cyclic-AMP responsive element CRE consensus oligo for EMSA #1.
XX
KW   Cyclic-AMP responsive element; electrophoretic mobility shift assay;
KW   EMSA; CRE; NF-kappaB; cancer; tetraethyl thiuram disulphide;
KW   dithiocarbamate; tumour; metal ion; copper ion; cytokine;
KW   ceruloplasmin; anticancer; cytostatic; ss.
XX
OS   Synthetic.
XX
PN   WO200117522-A1.
XX
PD   15-MAR-2001.
XX
PF   15-NOV-1999;   99WO-US27193.
XX
PR   08-SEP-1999;   99US-0392122.
XX
PA   (CHAR-) CHARLOTTE-MECKLENBURG HOSPITAL AUTHORITY.
XX
PI   Kennedy TP;
XX
DR   WPI; 2001-281426/29.
XX
PT   Treating cancer, e.g. melanoma, lung cancer, breast cancer or prostatic
PT   carcinoma, comprises administration of a thiuram disulfide optionally
PT   in combination with a heavy metal ion, ceruloplasmin or a cytokine e.g.
PT   interferon-alpha   -
XX
PS   Disclosure; Page 24; 60pp; English.
XX
CC   The present invention describes a method for treating established cancer
CC   in a mammal. The method comprises administering a thiuram disulfide (I).
CC   (I) has anticancer and cytostatic activities. (I) induces apoptosis and
CC   G2 cell-cycle arrest, by reducing expression of cyclin A (by preventing
CC   binding of transcription factors to DNA regulatory elements involved in
CC   control of cyclin A expression). The method can be used to treat cancers,
CC   e.g. melanoma, non-small cell lung cancer, small cell lung cancer, renal
CC   cancer, colorectal cancer, breast cancer, pancreatic cancer, gastric
CC   cancer, bladder cancer, ovarian cancer, uterine cancer, lymphoma and
CC   prostate cancer, especially melanoma, lung cancer, breast cancer and
CC   prostate carcinoma. The tumour-inhibiting effect of (I) is dependent on
CC   heavy metal ions, so administering (I) together with such ions (or with
CC   their intracellular carriers, e.g. ceruloplasmin or with serum
CC   ceruloplasmin stimulants, e.g. interferon-alpha) will increase the
CC   antiproliferative/antineoplastic affect. (I) also potentiates the
CC   effect of standard anticancer agents. (I) is already known for treating
CC   alcohol abuse (inhibitor of aldehyde dehydrogenase) and is relatively
CC   nontoxic and safe. The present sequence represents a cyclic-AMP
CC   responsive element CRE consensus oligonucleotide for use in an
CC   electrophoretic mobility shift assay (EMSA), which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 27 BP; 8 A; 5 C; 9 G; 5 T; 0 other;

  Query Match            100.0%;  Score 20;  DB 22;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23


RESULT 13
AAF87957/c
ID   AAF87957 standard; DNA; 27 BP.
XX
AC   AAF87957;
XX
DT   20-JUL-2001  (first entry)
XX
DE   Cyclic-AMP responsive element CRE consensus oligo for EMSA #2.
XX
```

```
KW   Cyclic-AMP responsive element; electrophoretic mobility shift assay;
KW   EMSA; CRE; NF-kappaB; cancer; tetraethyl thiuram disulphide;
KW   dithiocarbamate; tumour; metal ion; copper ion; cytokine;
KW   ceruloplasmin; anticancer; cytostatic; ss.
XX
OS   Synthetic.
XX
PN   WO200117522-A1.
XX
PD   15-MAR-2001.
XX
PF   15-NOV-1999;   99WO-US27193.
XX
PR   08-SEP-1999;   99US-0392122.
XX
PA   (CHAR-) CHARLOTTE-MECKLENBURG HOSPITAL AUTHORITY.
XX
PI   Kennedy TP;
XX
DR   WPI; 2001-281426/29.
XX
PT   Treating cancer, e.g. melanoma, lung cancer, breast cancer or prostatic
PT   carcinoma, comprises administration of a thiuram disulfide optionally
PT   in combination with a heavy metal ion, ceruloplasmin or a cytokine e.g.
PT   interferon-alpha -
XX
PS   Disclosure; Page 24; 60pp; English.
XX
CC   The present invention describes a method for treating established cancer
CC   in a mammal. The method comprises administering a thiuram disulfide (I).
CC   (I) has anticancer and cytostatic activities. (I) induces apoptosis and
CC   G2 cell-cycle arrest, by reducing expression of cyclin A (by preventing
CC   binding of transcription factors to DNA regulatory elements involved in
CC   control of cyclin A expression). The method can be used to treat cancers,
CC   e.g. melanoma, non-small cell lung cancer, small cell lung cancer, renal
CC   cancer, colorectal cancer, breast cancer, pancreatic cancer, gastric
CC   cancer, bladder cancer, ovarian cancer, uterine cancer, lymphoma and
CC   prostate cancer, especially melanoma, lung cancer, breast cancer and
CC   prostate carcinoma. The tumour-inhibiting effect of (I) is dependent on
CC   heavy metal ions, so administering (I) together with such ions (or with
CC   their intracellular carriers, e.g. ceruloplasmin or with serum
CC   ceruloplasmin stimulants, e.g. interferon-alpha) will increase the
CC   antiproliferative/antineoplastic affect. (I) also potentiates the
CC   effect of standard anticancer agents. (I) is already known for treating
CC   alcohol abuse (inhibitor of aldehyde dehydrogenase) and is relatively
CC   nontoxic and safe. The present sequence represents a cyclic-AMP
CC   responsive element CRE consensus oligonucleotide for use in an
CC   electrophoretic mobility shift assay (EMSA), which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 27 BP; 5 A; 9 C; 5 G; 8 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 22;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db         24 GATTGCCTGACGTCAGAGAG 5


RESULT 14
AAF76267
ID   AAF76267 standard; DNA; 27 BP.
XX
AC   AAF76267;
XX
DT   05-JUN-2001  (first entry)
XX
DE   cAMP response element (CRE) competitor EMSA probe.
XX
KW   NF-kappa-B; nuclear factor-kappa-B; cAMP response element; CRE;
KW   nuclear translocation inhibition; heparin; internalisation;
KW   NF-kappa-B dependent gene expression inhibition; cytokine;
KW   tumour necrosis factor; TNF; interleukin; IL-1; IL-2; IL-6; IL-8;
KW   interferon-beta; interferon-gamma; tissue factor-1; complement;
KW   inducible nitric oxide synthase; diabetic vascular disease;
KW   heart failure; asthma; sepsis; ischaemic-reperfusion injury;
KW   electrophoretic mobility shift assay; competitor EMSA probe; ds.
XX
OS   Unidentified.
XX
PN   WO200119376-A2.
XX
PD   22-MAR-2001.
XX
PF   12-SEP-2000; 2000WO-US24910.
XX
PR   13-SEP-1999;   99US-0395081.
XX
PA   (CHAR-) CHARLOTTE-MECKLENBURG HOSPITAL AUTHORITY.
XX
PI   Kennedy TP;
XX
DR   WPI; 2001-244698/25.
XX
PT   Inhibiting NF-kappa-B activity, useful for treating e.g. diabetic
PT   vascular disease, heart failure, asthma and sepsis, comprises
PT   administering heparin to cells in patient to inhibit translocation of
PT   NF-kappa-B from cytoplasm to nucleus -
XX
PS   Examples; Page 22; 68pp; English.
XX
CC   The invention relates to a method of inhibiting nuclear factor-kappa-B
CC   (NF-kappa-B) activity in a patient, comprising the administration of
CC   heparin to the cells in the patient, such that the heparin is
CC   internalised into the cytoplasm of cells in the patient. The invention
CC   is based on the discovery that heparin is able to block the
CC   translocation of NF-kappa-B from the cytoplasm to the nucleus. This in
CC   turn inhibits NF-kappa-B dependent gene expression. Such NF-kappa-B
CC   dependent genes include genes encoding cytokines such as tumour necrosis
CC   factor (TNF), IL-1 (interleukin-1), IL-2, IL-6, IL-8, interferon-beta,
CC   interferon-gamma, tissue factor-1, complement and inducible nitric
CC   oxide synthase. The method of the invention is used for treating or
CC   preventing diabetic vascular disease, heart failure, asthma, sepsis and
CC   ischaemic-reperfusion injury. Heparin may be administered in combination
CC   with other active agents that treat or prevent another disease or
CC   symptom in the patient, e.g., antiviral agents, antibiotics, antifungal
CC   agents and antiinflammatory agents. The method of the invention offers
CC   significant advantages over prior art treatments for the above
CC   conditions. Heparin is relatively non-toxic and safe, and should not
CC   produce the side effects such as hypertension, glucose intolerance
CC   and bone demineralisation that are encountered with the use of
CC   glucocorticoids for blocking the NF-kappa-B nuclear translocation.
CC   Additionally, heparin is readily available and easily used. Sequences
CC   AAF76266-AAF76267 represents EMSA (electrophoretic mobility shift assay)
CC   probes used to measure the effect of heparin on NF-kappa-B nuclear
CC   translocation. EMSA probe AAF76266 comprises a consensus NF-kappa-B
CC   response element, and EMSA competitor probe AAF76267 comprises a
CC   cAMP response element (CRE).
XX
SQ   Sequence 27 BP; 8 A; 5 C; 9 G; 5 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 22;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23


RESULT 15
ABA92274
ID   ABA92274 standard; DNA; 27 BP.
XX
```

AC ABA92274;
XX
DT 10-JUN-2002 (first entry)
XX
DE CRE binding site oligonucleotide, used in EMSA.
XX
KW CRE; electrophoretic mobility shift assay; EMSA; ds.
XX
OS Homo sapiens.
XX
PN WO200215912-A1.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26527.
XX
PR 25-AUG-2000; 2000US-228201P.
PR 26-OCT-2000; 2000US-243295P.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Ratan RR, Chatterjee S;
XX
DR WPI; 2002-242023/29.
XX
PT Diagnosing and treating diseases associated with oxidative stress, DNA
PT damage or growth factor depletion, e.g. Alzheimer's and Parkinson's, by
PT administering e.g. mithramycin, chromomycin, daunomycin, olivomycin and
PT WP631 -
XX
PS Example 1; Page 23; 69pp; English.
XX
CC The present sequence is that of a CRE binding site oligonucleotide,
CC which was radiolabelled and used in an electrophoretic mobility shift
CC assay (EMSA) to parallel an EMSA performed with Sp-1 oligonucleotides
CC (see ABA92271-72). The EMSAs were used to determine the effect of
CC oxidative stress on Sp-1 DNA binding, and the effects of candidate
CC compounds on Sp-1 protein levels. SP-1 DNA binding activity in
CC cortical neurons was shown to be low, but was dramatically enhanced
CC by oxidative stress. The invention provides methods for detecting
CC and treating diseases associated with oxidative stress, DNA damage
CC or growth factor depletion, and identifying agents for their
CC treatment. A compound is deemed to be an inhibitor of oxidative
CC stress, DNA damage, growth factor depletion or cell death if it
CC reduces the protein level of an Sp family member or if it decreases
CC the binding of an Sp family member to DNA. A method for preventing
CC or treating a disease or disorder of the nervous system, the ageing
CC process or associated with apoptosis involves administering a
CC compound that inhibits the induction of an Sp family member or the
CC binding of an Sp family member to DNA, e.g. mithramycin,
CC chromomycin, daunomycin, olivomycin or WP631. Diseases and
CC disorders that can be treated include Alzheimer's disease,
CC Creutzfeldt-Jacob disease, kuru, Huntington's disease, aneurysm,
CC stroke associated with an increase in blood pressure, spinal cord
CC disease, spinal cord injury, brain injury, multiple system atrophy,
CC amylotrophic lateral sclerosis, progressive supranuclear palsy,
CC neurodegeneration associated with the ageing process, mitochondrial
CC disease, HIV infection, herpes infection and multiple sclerosis
CC (all claimed).
XX
SQ Sequence 27 BP; 8 A; 5 C; 9 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GATTGCCTGACGTCAGAGAG 20
              ||||||||||||||||||||
Db          4 GATTGCCTGACGTCAGAGAG 23
```

Search completed: June 26, 2003, 12:16:10
Job time : 228.158 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 09:43:36 ; Search time 226.087 Seconds
(without alignments)
199.215 Million cell updates/sec

Title: US-09-355-254F-21
Perfect score: 20
Sequence: 1 gtatttcccagaaaaggaac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	17	AAT33591	FcgRI gene interle
2	20	100.0	20	17	AAT03680	Fc-gamma RI IL4 St
3	20	100.0	20	19	AAV46013	Immune adjuvant ST
4	20	100.0	20	20	AAX28360	Probe for Human St
5	20	100.0	20	20	AAX28365	PCR primer for Hum
6	20	100.0	20	21	AAZ92039	STAT5 binding sequ
7	20	100.0	20	24	AAL39188	Murine Toll-like r
8	20	100.0	42	15	AAQ63869	Protein binding mo
9	20	100.0	100	22	AAA54479	DNA fragment compr
c 10	20	100.0	100	22	AAA54480	DNA fragment compr
11	20	100.0	100	24	ABA93801	KZ134 plasmid cons
c 12	20	100.0	100	24	ABA93802	KZ134 plasmid cons
13	20	100.0	100	24	AAS20691	Plasmid KZ 134 oli
c 14	20	100.0	100	24	AAS20692	Plasmid KZ 134 oli
15	20	100.0	100	24	AAD22953	BaF3/KZ134/zalpha1
c 16	20	100.0	100	24	AAD22954	BaF3/KZ134/zalpha1
17	20	100.0	244	15	AAQ63438	Human Fc gamma R1
18	20	100.0	824	22	AAS05968	KZ159/mIL4 reporte
c 19	19	95.0	21710	22	AAS42185	Genomic sequence #
20	18.4	92.0	62	18	AAT76784	GAS motif biotinyl
21	18.4	92.0	655	24	ABK84058	Human cDNA differe
22	17.4	87.0	26	18	AAT59532	Fc(gamma)RI promot
23	17.4	87.0	149671	24	ABK84797	Human cDNA differe
24	17	85.0	17	17	AAT31284	STAT probe GRR. S
c 25	17	85.0	17	21	AAA28829	Oligo GRR, contain
c 26	17	85.0	17	24	AAD30634	Human GRR oligo #2
c 27	17	85.0	17	24	AAD27140	GRR probe lower st
28	16.8	84.0	1001	21	AAC57470	Arachidonic acid m
c 29	16.4	82.0	474	22	AAK62617	Human immune/haema
30	16.4	82.0	806	22	AAH05391	Human cDNA clone (
31	16.4	82.0	5749	23	ABL28262	Drosophila melanog
c 32	16.4	82.0	17070	22	AAK80632	Human immune/haema
33	16.4	82.0	27324	22	ABA89226	Escherichia coli p
c 34	16.4	82.0	36651	24	AAD28072	Human kinase genom
c 35	16.4	82.0	76798	24	ABN97454	Gene #3952 used to
36	16.4	82.0	130263	24	ABK83573	Human cDNA differe
37	16.4	82.0	1503900	22	AAK95240	Human neuregulin-1
38	16.4	82.0	1503900	22	AAK96733	Human neuregulin-1
39	16	80.0	25	16	AAQ89341	Fc-gamma-RI recept
40	16	80.0	30	17	AAT37047	Probe containing I
c 41	16	80.0	409	22	AAI91382	Human polynucleoti
c 42	16	80.0	428	22	AAK55159	Human immune/haema
c 43	15.8	79.0	239	21	AAC04878	Human secreted pro
44	15.8	79.0	375	22	AAI86131	Human polynucleoti
c 45	15.8	79.0	386	22	AAI80344	Human polynucleoti

ALIGNMENTS

RESULT 1
AAT33591
ID AAT33591 standard; DNA; 20 BP.
XX
AC AAT33591;
XX
DT 06-DEC-1996 (first entry)
XX
DE FcgRI gene interleukin-4 response element.
XX
KW Interleukin-4 response element; FcgRI; probe; Stat 5
KW signal transducer and activator of transcription 5; Stat 5;
KW interleukin-2; signal transduction; cell proliferation;
KW immune disorder; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1
FT /*tag= a
FT /note= "5' biotin label"
XX
PN WO9626292-A1.
XX
PD 29-AUG-1996.
XX
PF 22-FEB-1996; 96WO-US02457.
XX
PR 23-FEB-1995; 95US-0393333.
XX
PA (TULA-) TULARIK INC.
XX

PI Hou J, McKnight SL, Schindler U;
XX
DR WPI; 1996-402382/40.
XX
PT Human signal transducer and activator of transcription 5 protein -
PT used for treating cellular proliferation disorders, pref. of immune
PT cells
XX
PS Example; Page 22; 42pp; English.
XX
CC The interleukin-4 (IL-4) response element (AAT33591) of the gene
CC encoding FcgRIA was biotinylated and attached to streptavidin
CC agarose to form a DNA-affinity resin. Proteins obtd. from nuclear
CC extracts of IL-2-induced human lymphocyte YT cells were mixed with
CC the resin. After incubation, the affinity matrix was washed with
CC buffer or with buffer contg. a mutated variant (AAT33592) of the IL-4
CC response element to isolate IL-2-induced transcription factors.
CC Complexes between DNA and an isolated protein were visualised by a
CC gel mobility shift assay utilising the same IL-4 response elements
CC as probes.
XX
SQ Sequence 20 BP; 8 A; 4 C; 4 G; 4 T; 0 other;

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 1 GTATTTCCCAGAAAAGGAAC 20

RESULT 2
AAT03680
ID AAT03680 standard; cDNA; 20 BP.
XX
AC AAT03680;
XX
DT 29-MAR-1996 (first entry)
XX
DE Fc-gamma RI IL4 Stat cytokine binding domain cDNA.
XX
KW Interleukin-4 signal transducer and activator of transcription;
KW IL-4 Stat; transcription factor; immunosuppressive; Fc-gamma RI; ds.
XX
OS Homo sapiens.
XX
PN EP692488-A2.
XX
PD 17-JAN-1996.
XX
PF 05-JUL-1995; 95EP-0304715.
XX
PR 15-JUL-1994; 94US-0276099.
PR 05-JUL-1994; 94US-0269604.
XX
PA (TULA-) TULARIK INC.
XX
PI Hou J, McKnight SL;
XX
DR WPI; 1996-070143/08.
XX
PT IL-4 signal transducer and activator of transcription (IL-4 Stat)
PT peptide(s) - bind to natural intracellular IL-4 Stat binding target
PT and are useful to identify cpds. for treatment and diagnosis of
PT immune diseases
XX
PS Disclosure; Page 5; 22pp; English.
XX
CC Interleukin-4 signal transducer and activator of transcription
CC (IL-4 Stat) (AAR88320) is characterized by selective binding to
CC intracellular domains of cytokine receptors; nucleic acids
CC encoding IL-4 Stat binding sites are given in AAT03680-86.
CC Preferred binding sites include 2 trinucleotides (TTC and GAA)
CC separated by 1-5 nucleotides.
XX
SQ Sequence 20 BP; 8 A; 4 C; 4 G; 4 T; 0 other;

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 1 GTATTTCCCAGAAAAGGAAC 20

RESULT 3
AAV46013
ID AAV46013 standard; DNA; 20 BP.
XX
AC AAV46013;
XX
DT 16-OCT-1998 (first entry)
XX
DE Immune adjuvant STAT5/6.
XX
KW Immune system; adjuvant; vaccine; cancer; prophylactic; pathogenicity;
KW modulator; tolerance; regulator; helper cell; antigen; immunoglobulin;
KW Ig class; autoimmune response; T-cell; B-cell; tumour; ss.
XX
OS Class Bacteria.
XX
PN EP855184-A1.
XX
PD 29-JUL-1998.
XX
PF 23-JAN-1997; 97EP-0101019.
XX
PR 23-JAN-1997; 97EP-0101019.
XX
PA (HEEG/) HEEG K.
PA (LIPF/) LIPFORD G B.
PA (WAGN/) WAGNER H.
XX
PI Heeg K, Lipford GB, Wagner H;
XX
DR WPI; 1998-389630/34.
XX
PT Antigenic composition comprises polynucleotide fragment and antigen
PT - used as vaccine to treat or prevent e.g. cancer or pathogen
PT infections and to modulate immune response e.g. tolerance break and
PT regulation of TH1/TH2 cells
XX
PS Example 5; Page 9; 28pp; English.
XX
CC AAV45993-V46019 are fragments of bacterial polynucleotides which are
CC used as immune adjuvants for inclusion into vaccines to treat cancer and
CC for prophylaxis and/or treatment of conditions caused by pathogenic
CC micro-organisms. The polynucleotide is used for modulation of an immune
CC response and the modulation is selected from the group break of
CC tolerance, regulation of TH1/TH2 helper cell responses, switch of Ig
CC classes, treatment of autoimmune responses and induction of tolerances.
CC DNA oligomers are used to enhance the reactivity of immune cells to
CC viral, bacterial and parasitic antigens, to break tolerance in anergic T
CC and B cells e.g. against tumour antigens, as adjuvants in vaccination
CC against tumour-defined antigens and immunostimulatory substances in an
CC immune response against tumours and to suppress immune reactions of the
CC innate and acquired immune system. The composition is inexpensive and
CC stable and does not cause lethal shock, which happens with prior art
CC bacterial sequences.
XX
SQ Sequence 20 BP; 8 A; 4 C; 4 G; 4 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1;

```
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 4
AAX28360
ID   AAX28360 standard; DNA; 20 BP.
XX
AC   AAX28360;
XX
DT   18-JUN-1999  (first entry)
XX
DE   Probe for Human Stat6 coding sequence.
XX
KW   Stat6; Stat6b; human; signal transducers and activators of transcription;
KW   isoform; myeloid cancer; asthma; sarcoma; scleroderma; fibrotic disease;
KW   bone marrow fibrosis; AIDS; Stat6c; probe; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9910493-A1.
XX
PD   04-MAR-1999.
XX
PF   27-AUG-1998;   98WO-US17821.
XX
PR   05-JAN-1998;   98US-0070397.
PR   27-AUG-1997;   97US-0056075.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Larochelle WJ,  Patel B,  Pierce JH;
XX
DR   WPI; 1999-214517/18.
XX
PT   New isoforms of Stat6  - having differential effects on the
PT   modulation of Stat6 activity in cells
XX
PS   Example 1; Page 20; 88pp; English.
XX
CC   This sequence is a probe for DNA encoding human Stat6 (signal transducers
CC   and activators of transcription). The invention relates to
CC   isoforms of human Stat6. The detection and quantitation of DNA
CC   or mRNA encoding Stat6 and/or Stat6b and/or Stat6c can be used to detect
CC   differential expression of Stat6 isoforms in numerous diseases, including
CC   myeloid cancer, asthma, sarcoma, scleroderma, bone marrow fibrosis,
CC   fibrotic diseases and AIDS. The nucleic acids can be used to screen
CC   genomic or cDNA libraries or to identify complementary sequences. The
CC   identification of the genetic locus of the Stat6 gene can be used for
CC   detection of chromosomal aberrations and translocations involving the
CC   Stat6 gene. Antibodies against the isoforms can be used to detect the
CC   presence of Stat6 and/or Stat6b and/or Stat6c in a sample. Because of the
CC   variation of the roles of Stat6b and Stat6c in regulating gene
CC   transcription the isolated and purified forms can be used to study gene
CC   regulation and in screening assays for identifying drug candidates which
CC   may be used as agonists or antagonists. The two polypeptides may also be
CC   used in gene therapy protocols. In particular, Stat6b and/or Stat6c can
CC   therapeutically modulate the development and differentiation of B and
CC   T cells and can enhance IL-4 immunological function in immunocompromised
CC   individuals. Stat6 activation correlates with functional responses
CC   induced by interleukin-4 (IL-4), IL-13 and platelet-derived growth factor
CC   (PDGF). Stat6b when compared Stat6 is an attenuated regulator of gene
CC   transcription. Stat6c is a dominant negative regulator of gene
CC   transcription.
XX
SQ   Sequence 20 BP; 8 A; 4 C; 4 G; 4 T; 0 other;

  Query Match           100.0%;  Score 20;  DB 20;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 4.1;

  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 5
AAX28365
ID   AAX28365 standard; DNA; 20 BP.
XX
AC   AAX28365;
XX
DT   18-JUN-1999  (first entry)
XX
DE   PCR primer for Human Stat6 coding sequence.
XX
KW   Stat6; Stat6b; human; signal transducers and activators of transcription;
KW   isoform; myeloid cancer; asthma; sarcoma; scleroderma; fibrotic disease;
KW   bone marrow fibrosis; AIDS; Stat6c; primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9910493-A1.
XX
PD   04-MAR-1999.
XX
PF   27-AUG-1998;   98WO-US17821.
XX
PR   05-JAN-1998;   98US-0070397.
PR   27-AUG-1997;   97US-0056075.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Larochelle WJ,  Patel B,  Pierce JH;
XX
DR   WPI; 1999-214517/18.
XX
PT   New isoforms of Stat6  - having differential effects on the
PT   modulation of Stat6 activity in cells.
XX
PS   Example 2; Page 30; 88pp; English.
XX
CC   This sequence is a primer for DNA encoding human Stat6 (signal
CC   transducers and activators of transcription). The invention relates to
CC   isoforms of human Stat6. The detection and quantitation of DNA
CC   or mRNA encoding Stat6 and/or Stat6b and/or Stat6c can be used to detect
CC   differential expression of Stat6 isoforms in numerous diseases, including
CC   myeloid cancer, asthma, sarcoma, scleroderma, bone marrow fibrosis,
CC   fibrotic diseases and AIDS. The nucleic acids can be used to screen
CC   genomic or cDNA libraries or to identify complementary sequences. The
CC   identification of the genetic locus of the Stat6 gene can be used for
CC   detection of chromosomal aberrations and translocations involving the
CC   Stat6 gene. Antibodies against the isoforms can be used to detect the
CC   presence of Stat6 and/or Stat6b and/or Stat6c in a sample. Because of the
CC   variation of the roles of Stat6b and Stat6c in regulating gene
CC   transcription the isolated and purified forms can be used to study gene
CC   regulation and in screening assays for identifying drug candidates which
CC   may be used as agonists or antagonists. The two polypeptides may also be
CC   used in gene therapy protocols. In particular, Stat6b and/or Stat6c can
CC   therapeutically modulate the development and differentiation of B and
CC   T cells and can enhance IL-4 immunological function in immunocompromised
CC   individuals. Stat6 activation correlates with functional responses
CC   induced by interleukin-4 (IL-4), IL-13 and platelet-derived growth factor
CC   (PDGF). Stat6b when compared Stat6 is an attenuated regulator of gene
CC   transcription. Stat6c is a dominant negative regulator of gene
CC   transcription.
XX
SQ   Sequence 20 BP; 8 A; 4 C; 4 G; 4 T; 0 other;

  Query Match           100.0%;  Score 20;  DB 20;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 4.1;
```

```
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 6
AAZ92039
ID   AAZ92039 standard; DNA; 20 BP.
XX
AC   AAZ92039;
XX
DT   08-JUN-2000  (first entry)
XX
DE   STAT5 binding sequence oligonucleotide STAT5/STAT6.
XX
KW   STAT5 protein; signal transducer and activator of transcription 5;
KW   protein binding sequence; transcription factor modulator; inhibitor;
KW   malignant cell removal; proliferative malignancy; neoplastic disease;
KW   immunological disorder; inflammatory disorder; therapy; ds.
XX
OS   Synthetic.
XX
PN   WO200006696-A2.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-US17366.
XX
PR   30-JUL-1998;   98US-0094695.
XX
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Zuckerman KS,  Liu RY;
XX
DR   WPI; 2000-195281/17.
XX
PT   Therapeutic agent for treating transcription factor-related illnesses
PT   such as proliferative malignancies, comprises an oligonucleotide for
PT   regulating transcription factor function -
XX
PS   Claim 15; Page 34; 43pp; English.
XX
CC   This sequence represents a STAT5 (signal transducer and activator of
CC   transcription 5) protein binding sequence. The invention relates to a
CC   therapeutic agent comprising an effective amount of an oligonucleotide
CC   (I) for modulating the function of transcription factors and a
CC   pharmaceutical acceptable carrier. The oligonucleotides can be used in a
CC   method of removing malignant cells in vitro. The oligonucleotides can be
CC   used in compositions to inhibit transcription factors in illnesses where
CC   transcription factors play a role, especially proliferative malignancies,
CC   neoplastic diseases, and immunological and inflammatory disorders.
XX
SQ   Sequence 20 BP; 8 A; 4 C; 4 G; 4 T; 0 other;

  Query Match            100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 4.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 7
AAL39188
ID   AAL39188 standard; DNA; 20 BP.
XX
AC   AAL39188;
XX
DT   05-SEP-2002  (first entry)
XX
DE   Murine Toll-like receptor related CpG DNA SEQ ID No 63.
XX
KW   Murine Toll-like receptor; TLR9; TLR7; TLR8; ISNA; ds.
XX
OS   Unidentified.
XX
PN   WO200222809-A2.
XX
PD   21-MAR-2002.
XX
PF   17-SEP-2001; 2001WO-US29229.
XX
PR   15-SEP-2000; 2000US-233035P.
PR   23-JAN-2001; 2001US-263657P.
PR   17-MAY-2001; 2001US-291726P.
PR   22-JUN-2001; 2001US-300210P.
XX
PA   (COLE-) COLEY PHARM GMBH.
XX
PI   Bauer S,  Lipford G,  Wagner H;
XX
DR   WPI; 2002-393964/42.
XX
PT   New isolated murine Toll-like receptor (TLR)9, TLR7, TLR8 polypeptides,
PT   useful for identifying species specificity of immunostimulatory nucleic
PT   acid and identifying immunostimulatory nucleic acids -
XX
PS   Disclosure; Page 76; 195pp; English.
XX
CC   The invention relates to isolated murine Toll-like receptors (TLR)9,
CC   TLR7 and TLR8 polypeptides. These polypeptides comprise fully defined
CC   sequences of 1032, 1050 or 1032 amino acids as given in specification, or
CC   their fragments, where TLR9, TLR7 and TLR8 polypeptides or their
CC   fragments have an amino acid sequence which is identical to human TLR9,
CC   TLR7 or TLR8 polypeptides or their fragment except for at least one amino
CC   acid of a murine TLR polypeptide. The isolated nucleic acids of the
CC   invention are useful for inhibiting TLR9 signalling activity in a cell.
CC   TLR7, TLR8 and TLR9 polypeptides are useful for identifying nucleic acid
CC   molecules which interact with a TLR polypeptide or its fragment. The
CC   TLR7, TLR8 or TLR9 polypeptides are also useful for identifying ISNA. The
CC   TLR7, TLR8 and TLR9 polypeptides are also useful for comparing TLR9
CC   signalling activity of a test compound (that is not a nucleic acid, and
CC   is a polypeptide or a part of a combinatorial library of compounds) with
CC   an ISNA. The TLR7, TLR8 and TLR9 polypeptides are also useful for
CC   identifying species specificity of an ISNA. The isolated nucleic acids of
CC   the invention are useful as probes or primers. This polynucleotide
CC   sequence represents DNA relating to the isolated Toll-like receptors of
CC   the invention.
XX
SQ   Sequence 20 BP; 8 A; 4 C; 4 G; 4 T; 0 other;

  Query Match            100.0%;  Score 20;  DB 24;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 4.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTATTTCCCAGAAAAGGAAC 20
              ||||||||||||||||||||
Db          1 GTATTTCCCAGAAAAGGAAC 20


RESULT 8
AAQ63869
ID   AAQ63869 standard; DNA; 42 BP.
XX
AC   AAQ63869;
XX
DT   10-NOV-1994  (first entry)
XX
DE   Protein binding motif GIRE from human R gene 5'-UTR.
XX
KW   Immunoglobulin; IgG receptor; gamma-interferon activation; g-IFN;
KW   haematopoietic cell; protein factor; binding site; GRR motif;
KW   human high affinity Fc receptor for IgG; human Fc gamma R1; R gene;
```

KW DNA response element; MATE motif; GIRE motif; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN FR2696181-A.
XX
PD 01-APR-1994.
XX
PF 25-SEP-1992; 92FR-0011498.
XX
PR 25-SEP-1992; 92FR-0011498.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Benech P, Perez C, Wietzerbin J;
XX
DR WPI; 1994-128281/16.
XX
PT New nucleotide sequences are specific target for proteins
PT exclusive to haematopoietic cells - used to impart, to partic.
PT cells, a response to gamma-interferon, e.g. for gene therapy
XX
PS Claim 10; Page 25; 37pp; French.
XX
CC The region immediately upstream of the sequence which codes for the
CC human high affinity Fc receptor for IgG (the R gene) comprises
CC binding sites specific for haematopoietic cells (see AAQ63866-8).
CC The sequence AAQ63869, designated the "GIRE" motif, is recognised by
CC proteins present in both haematopoietic and non-haematopoietic cells
CC and controls induction by interferon-gamma. The GIRE motif is used
CC (opt. in multimeric form) with AAQ63866-8 for conferring on a gene
CC a transcriptional activity limited to haematopoietic cells. The DNA
CC motif thus has potential in gene therapy.
XX
SQ Sequence 42 BP; 14 A; 6 C; 10 G; 12 T; 0 other;

Query Match 100.0%; Score 20; DB 15; Length 42;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 23 GTATTTCCCAGAAAAGGAAC 42

RESULT 9
AAA54479
ID AAA54479 standard; DNA; 100 BP.
XX
AC AAA54479;
XX
DT 11-APR-2001 (first entry)
XX
DE DNA fragment comprising STAT transcription factors.
XX
KW zcytor 10 cytokine receptor; cytokine; receptor; antibody; ligand;
KW binding; detection; modulation; recombinant cell;
KW haematopoietic cell; lymphoid cell; myeloid cell; lymph;
KW immune system; blood; bone; inflammatory response; inflammation;
KW spleen; human; primer; ss.
XX
OS Synthetic.
XX
PN WO200068381-A1.
XX
PD 16-NOV-2000.
XX
PF 11-MAY-2000; 2000WO-US12924.
XX
PR 11-MAY-1999; 99US-0309861.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Presnell SR, Foster DC, Hammond AK, Lok S;
XX
DR WPI; 2001-016096/02.
XX
PT New cytokine receptor mouse zcytor 10, useful for detecting ligands
PT that stimulate proliferation or development of haematopoietic,
PT lymphoid and myeloid cells
XX
PS Example 19; Page 128; 134pp; English.
XX
CC Isolating a nucleotide which encodes the zcytor 10 cytokine
CC receptor enables the production of recombinant cells expressing the
CC receptor. Those cells can then be used to detect the presence of a
CC modulator of zcytor10 protein by culturing the cells in the presence
CC of a test ligand and comparing levels of activity of mouse zcytor10
CC in the presence and absence of the test sample. Similarly, detection
CC of zcytor10 receptor ligand within a test sample can be achieved.
CC The method comprising contacting a test sample containing an amino
CC acid sequence from Cys15 or Gly25 to Pro230 of the zcytor 10
CC cytokine receptor and detecting the binding of the polypeptide to a
CC ligand in the sample. Specified peptide fragments of the zcytor 10
CC cytokine receptor and the methods described are used to identify
CC ligands that stimulate the proliferation and/or development of
CC haematopoietic, lymphoid and myeloid cells. Peptide fragments of
CC the cytokine receptor are useful for treating lymphoid, immune,
CC inflammatory, splenic, blood or bone disorders and for generating
CC antibodies directed against the receptor. An exemplary luciferase
CC mammalian expression vector is the KZ134 plasmid which was
CC constructed with two complementary oligonucleotides (AAA54479,
CC AAA54480) which comprise STAT transcription factors from 4 genes
CC (a modified c-fos Sis element, the p21 SIE1 from the p21 WAF1
CC gene, the mammary gland response element of the Beta-casein gene
CC and a STAT inducible element of the Fcg R1 gene.
XX
SQ Sequence 100 BP; 24 A; 33 C; 17 G; 26 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 44 GTATTTCCCAGAAAAGGAAC 63

RESULT 10
AAA54480/c
ID AAA54480 standard; DNA; 100 BP.
XX
AC AAA54480;
XX
DT 11-APR-2001 (first entry)
XX
DE DNA fragment comprising STAT transcription factors.
XX
KW zcytor 10 cytokine receptor; cytokine; receptor; antibody; ligand;
KW binding; detection; modulation; recombinant cell;
KW haematopoietic cell; lymphoid cell; myeloid cell; lymph;
KW immune system; blood; bone; inflammatory response; inflammation;
KW spleen; human; primer; ss.
XX
OS Synthetic.
XX
PN WO200068381-A1.
XX
PD 16-NOV-2000.
XX
PF 11-MAY-2000; 2000WO-US12924.
XX
PR 11-MAY-1999; 99US-0309861.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX

PI Presnell SR, Foster DC, Hammond AK, Lok S;
XX
DR WPI; 2001-016096/02.
XX
PT New cytokine receptor mouse zcytor 10, useful for detecting ligands
PT that stimulate proliferation or development of haematopoietic,
PT lymphoid and myeloid cells
XX
PS Example 19; Page 128; 134pp; English.
XX
CC Isolating a nucleotide which encodes the zcytor 10 cytokine
CC receptor enables the production of recombinant cells expressing the
CC receptor. Those cells can then be used to detect the presence of a
CC modulator of zcytor10 protein by culturing the cells in the presence
CC of a test ligand and comparing levels of activity of mouse zcytor10
CC in the presence and absence of the test sample. Similarly, detection
CC of zcytor10 receptor ligand within a test sample can be achieved.
CC The method comprising contacting a test sample containing an amino
CC acid sequence from Cys15 or Gly25 to Pro230 of the zcytor 10
CC cytokine receptor and detecting the binding of the polypeptide to a
CC ligand in the sample. Specified peptide fragments of the zcytor 10
CC cytokine receptor and the methods described are used to identify
CC ligands that stimulate the proliferation and/or development of
CC haematopoietic, lymphoid and myeloid cells. Peptide fragments of
CC the cytokine receptor are useful for treating lymphoid, immune,
CC inflammatory, splenic, blood or bone disorders and for generating
CC antibodies directed against the receptor. An exemplary luciferase
CC mammalian expression vector is the KZ134 plasmid which was
CC constructed with two complementary oligonucleotides (AAA54479,
CC AAA54480) which comprise STAT transcription factors from 4 genes
CC (a modified c-fos Sis element, the p21 SIE1 from the p21 WAF1
CC gene, the mammary gland response element of the Beta-casein gene
CC and a STAT inducible element of the Fcg R1 gene.
XX
SQ Sequence 100 BP; 26 A; 17 C; 33 G; 24 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 61 GTATTTCCCAGAAAAGGAAC 42

RESULT 11
ABA93801
ID ABA93801 standard; DNA; 100 BP.
XX
AC ABA93801;
XX
DT 01-MAY-2002 (first entry)
XX
DE KZ134 plasmid construction oligonucleotide SEQ ID NO:43.
XX
KW Zcytor17; chromosome 5; 5q11; cytokine receptor; immunomodulatory;
KW antiinflammatory; antiviral; antirheumatic; antiarthritic; cytostatic;
KW muscular; lymphoid; immune; inflammatory; spleenic; blood; bone;
KW infection; immunosuppression; cytotoxicity; leukopenia; Crohn's disease;
KW autoimmune disease; rheumatoid arthritis; multiple sclerosis; cancer;
KW inflammatory disease; pancreatitis; inflammatory bowel disease;
KW PCR primer; probe; ss.
XX
OS Synthetic.
XX
PN WO200200721-A2.
XX
PD 03-JAN-2002.
XX
PF 26-JUN-2001; 2001WO-US20484.
XX
PR 26-JUN-2000; 2000US-214282P.
PR 29-JUN-2000; 2000US-214955P.
PR 08-FEB-2001; 2001US-267963P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sprecher CA, Presnell SR, Gao Z, Whitmore TE, Kuijper JL;
PI Maurer MF;
XX
DR WPI; 2002-090519/12.
XX
PT Isolated polynucleotide encoding a cytokine receptor zcytor17 which is
PT useful for treating and diagnosing lymphoid, immune, inflammatory,
PT spleenic, blood or bone disorders -
XX
PS Example 19; Page 190; 235pp; English.
XX
CC The present invention describes a cytokine receptor designated zcytor17.
CC Zcytor17 has immunomodulatory, antiinflammatory, antiviral, cytostatic,
CC antirheumatic, antiarthritic and muscular activities. The zcytor17
CC proteins are useful for treating and diagnosing lymphoid, immune,
CC inflammatory, spleenic, blood or bone disorders. Agonists or
CC anti-zcytor17 antibodies are useful in stimulating cell-mediated
CC immunity and for stimulating lymphocyte proliferation, such as in the
CC treatment of infections involving immunosuppression, including certain
CC viral infections. They are also useful for inducting cytotoxicity and
CC for treating leukopenias. Antagonist of zcytor17 polypeptides are useful
CC for treating autoimmune diseases (e.g. rheumatoid arthritis and multiple
CC sclerosis), inflammatory diseases (e.g. Crohn's disease), cancer,
CC pancreatitis, and inflammatory bowel disease. Zcytor17 was mapped to
CC chromosome 5, specifically to the 5q11 chromosomal region. ABA93767 to
CC ABA93843 and ABB05730 to ABB05745 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 100 BP; 24 A; 33 C; 17 G; 26 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 44 GTATTTCCCAGAAAAGGAAC 63

RESULT 12
ABA93802/c
ID ABA93802 standard; DNA; 100 BP.
XX
AC ABA93802;
XX
DT 01-MAY-2002 (first entry)
XX
DE KZ134 plasmid construction oligonucleotide SEQ ID NO:44.
XX
KW Zcytor17; chromosome 5; 5q11; cytokine receptor; immunomodulatory;
KW antiinflammatory; antiviral; antirheumatic; antiarthritic; cytostatic;
KW muscular; lymphoid; immune; inflammatory; spleenic; blood; bone;
KW infection; immunosuppression; cytotoxicity; leukopenia; Crohn's disease;
KW autoimmune disease; rheumatoid arthritis; multiple sclerosis; cancer;
KW inflammatory disease; pancreatitis; inflammatory bowel disease;
KW PCR primer; probe; ss.
XX
OS Synthetic.
XX
PN WO200200721-A2.
XX
PD 03-JAN-2002.
XX
PF 26-JUN-2001; 2001WO-US20484.
XX
PR 26-JUN-2000; 2000US-214282P.
PR 29-JUN-2000; 2000US-214955P.
PR 08-FEB-2001; 2001US-267963P.
XX

PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sprecher CA, Presnell SR, Gao Z, Whitmore TE, Kuijper JL;
PI Maurer MF;
XX
DR WPI; 2002-090519/12.
XX
PT Isolated polynucleotide encoding a cytokine receptor zcytor17 which is
PT useful for treating and diagnosing lymphoid, immune, inflammatory,
PT spleenic, blood or bone disorders -
XX
PS Example 19; Page 190; 235pp; English.
XX
CC The present invention describes a cytokine receptor designated zcytor17.
CC Zcytor17 has immunomodulatory, antiinflammatory, antiviral, cytostatic,
CC antirheumatic, antiarthritic and muscular activities. The zcytor17
CC proteins are useful for treating and diagnosing lymphoid, immune,
CC inflammatory, spleenic, blood or bone disorders. Agonists or
CC anti-zcytor17 antibodies are useful in stimulating cell-mediated
CC immunity and for stimulating lymphocyte proliferation, such as in the
CC treatment of infections involving immunosuppression, including certain
CC viral infections. They are also useful for inducting cytotoxicity and
CC for treating leukopenias. Antagonist of zcytor17 polypeptides are useful
CC for treating autoimmune diseases (e.g. rheumatoid arthritis and multiple
CC sclerosis), inflammatory diseases (e.g. Crohn's disease), cancer,
CC pancreatitis, and inflammatory bowel disease. Zcytor17 was mapped to
CC chromosome 5, specifically to the 5q11 chromosomal region. ABA93767 to
CC ABA93843 and ABB05730 to ABB05745 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 100 BP; 26 A; 17 C; 33 G; 24 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 61 GTATTTCCCAGAAAAGGAAC 42

RESULT 13
AAS20691
ID AAS20691 standard; DNA; 100 BP.
XX
AC AAS20691;
XX
DT 09-APR-2002 (first entry)
XX
DE Plasmid KZ 134 oligonucleotide ZC12749.
XX
KW Cytokine; zalphal1 Ligand; zalphal1 receptor; NK cell progenitor;
KW natural killer cell proliferation; T-cell proliferation;
KW B-cell proliferation; anti-tumour response; immune system;
KW immunostimulant; cytostatic; primer; ss.
XX
OS Synthetic.
XX
PN US6307024-B1.
XX
PD 23-OCT-2001.
XX
PF 09-MAR-2000; 2000US-0522217.
XX
PR 09-MAR-1999; 99US-123547P.
PR 11-MAR-1999; 99US-123904P.
PR 01-JUL-1999; 99US-142013P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
DR WPI; 2002-040208/05.
XX
PT New zalphal1 ligand polypeptides and polynucleotides, useful for
PT stimulating proliferation, activation, differentiation and/or induction
PT of inhibition of specialized cell function, or for stimulating an
PT antigenic response -
XX
PS Example 20; Column 149-150; 105pp; English.
XX
CC The present invention relates to the isolation of a novel cytokine,
CC zalphal1 Ligand and the polynucleotide encoding it. The invention
CC also gives the sequence for the zalphal1 receptor and the polynucleotide
CC encoding it. The zalphal1 Ligand polypeptide stimulates proliferation of
CC natural killer (NK) cells or NK cell progenitors, the activation of NK
CC cells, proliferation of T-cells, proliferation of B-cells stimulated
CC with anti-CD40 antibodies, stimulates an antigenic response in a mammal,
CC and reduces proliferation of B-cells stimulated with anti-IgM antibodies.
CC The zalphal1 Ligand polypeptide is also useful in preparing antibodies
CC that bind to zalphal1 Ligand epitopes. The zalphal1 Ligand
CC polynucleotides can be used as probes or primers to clone regions
CC of a zalphal1 Ligand gene, and in gene therapy. Zalphal1 Ligand may
CC also be used to identify inhibitors of its activity, to enhance the
CC generation of anti-tumour responses with or without the infusion of
CC donor lymphocytes, and to activate or stimulate the immune system.
CC The present sequence represents an oligonucleotide used to construct
CC plasmid KZ 134 in the methods of the present invention.
XX
SQ Sequence 100 BP; 25 A; 32 C; 17 G; 26 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 44 GTATTTCCCAGAAAAGGAAC 63

RESULT 14
AAS20692/c
ID AAS20692 standard; DNA; 100 BP.
XX
AC AAS20692;
XX
DT 09-APR-2002 (first entry)
XX
DE Plasmid KZ 134 oligonucleotide ZC12748.
XX
KW Cytokine; zalphal1 Ligand; zalphal1 receptor; NK cell progenitor;
KW natural killer cell proliferation; T-cell proliferation;
KW B-cell proliferation; anti-tumour response; immune system;
KW immunostimulant; cytostatic; primer; ss.
XX
OS Synthetic.
XX
PN US6307024-B1.
XX
PD 23-OCT-2001.
XX
PF 09-MAR-2000; 2000US-0522217.
XX
PR 09-MAR-1999; 99US-123547P.
PR 11-MAR-1999; 99US-123904P.
PR 01-JUL-1999; 99US-142013P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
DR WPI; 2002-040208/05.
XX
PT New zalphal1 ligand polypeptides and polynucleotides, useful for

PT stimulating proliferation, activation, differentiation and/or induction
PT of inhibition of specialized cell function, or for stimulating an
PT antigenic response -
XX
PS Example 20; Column 149-151; 105pp; English.
XX
CC The present invention relates to the isolation of a novel cytokine,
CC zalpha11 Ligand and the polynucleotide encoding it. The invention
CC also gives the sequence for the zalpha11 receptor and the polynucleotide
CC encoding it. The zalpha11 Ligand polypeptide stimulates proliferation of
CC natural killer (NK) cells or NK cell progenitors, the activation of NK
CC cells, proliferation of T-cells, proliferation of B-cells stimulated
CC with anti-CD40 antibodies, stimulates an antigenic response in a mammal,
CC and reduces proliferation of B-cells stimulated with anti-IgM antibodies.
CC The zalpha11 Ligand polypeptide is also useful in preparing antibodies
CC that bind to zalpha11 Ligand epitopes. The zalpha11 Ligand
CC polynucleotides can be used as probes or primers to clone regions
CC of a zalpha11 Ligand gene, and in gene therapy. Zalpha11 Ligand may
CC also be used to identify inhibitors of its activity, to enhance the
CC generation of anti-tumour responses with or without the infusion of
CC donor lymphocytes, and to activate or stimulate the immune system.
CC The present sequence represents an oligonucleotide used to construct
CC plasmid KZ 134 in the methods of the present invention.
XX
SQ Sequence 100 BP; 26 A; 17 C; 32 G; 25 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 61 GTATTTCCCAGAAAAGGAAC 42

RESULT 15
AAD22953
ID AAD22953 standard; DNA; 100 BP.
XX
AC AAD22953;
XX
DT 26-FEB-2002 (first entry)
XX
DE BaF3/KZ134/zalpha11 cell line constructing ZC12,749 oligonucleotide.
XX
KW Zalpha11; cytokine receptor; immunosuppressive; cytostatic; haemostatic;
KW inflammatory disorder; cell proliferation; immune disorder; cancer; SLE;
KW systemic lupus erythematosus; myasthenia gravis; rheumatoid arthritis;
KW diabetes; autoimmune disease; multiple sclerosis; ulcerative colitis;
KW inflammatory bowel disease; sepsis; Crohn's disease; viral infection;
KW asthma; ss.
XX
OS Unidentified.
XX
PN WO200177171-A2.
XX
PD 18-OCT-2001.
XX
PF 03-APR-2001; 2001WO-US10872.
XX
PR 05-APR-2000; 2000US-194731P.
PR 28-JUL-2000; 2000US-222121P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Sprecher CA, Novak JE, West JW, Presnell SR, Holly RD, Nelson AJ;
XX
DR WPI; 2002-025898/03.
XX
PT Novel soluble receptor polypeptides and polynucleotides used as
PT cytokine antagonist for stimulating ligand activity-induced
PT proliferation of hematopoietic cells and for suppressing immune
PT response in a mammal -
XX
PS Example 19; Page 213; 243pp; English.
XX
CC The invention relates to an isolated soluble zalpha11 cytokine receptor
CC polypeptide and their cDNA molecules. Zalpha proteins are useful for
CC inhibiting or antagonising the ligand activity-induced proliferation of
CC haematopoietic cells and haematopoietic cell progenitors preferably
CC lymphoid cells which are natural killer cells or cytotoxic T cells.
CC Zalpha is useful for treating immune and inflammatory disorders, for
CC reducing proliferation of neoplastic B or T cells, for suppressing an
CC immune response in a mammal exposed to an antigen or pathogen. Zalpha is
CC useful for treating diseases that require immune regulation including
CC autoimmune diseases such as rheumatoid arthritis, multiple sclerosis,
CC myasthenia gravis, systemic lupus erythematosus (SLE) and diabetes;
CC asthma, ulcerative colitis, inflammatory bowel disease, Crohn's disease,
CC sepsis, viral infection (dengue virus infection) and cancer. The present
CC sequence is an oligonucleotide used for BaF3/KZ134/zalpha11 cell line
CC construction.
XX
SQ Sequence 100 BP; 24 A; 33 C; 17 G; 26 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 100;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATTTCCCAGAAAAGGAAC 20
||||||||||||||||||||
Db 44 GTATTTCCCAGAAAAGGAAC 63

Search completed: June 26, 2003, 12:16:35
Job time : 229.158 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 09:56:18 ; Search time 1028.48 Seconds
(without alignments)
565.939 Million cell updates/sec

Title: US-09-355-254F-16
Perfect score: 20
Sequence: 1 gtccatttcccgtaaatctt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	6	A89795	A89795 Sequence 17
2		20	100.0	20	6	A90882	A90882 Sequence 17
3		20	100.0	20	6	AX455587	AX455587 Sequence
4		19	95.0	123291	9	AC104070	AC104070 Homo sapi
5		17.4	87.0	22	6	AX040434	AX040434 Sequence
6	c	17.4	87.0	105428	2	AC094579	AC094579 Rattus no
7	c	17.4	87.0	164383	2	AC127444	AC127444 Rattus no
8		17	85.0	19	6	I81948	I81948 Sequence 46
9	c	16.8	84.0	2274	1	AF498313	AF498313 Streptoco
10	c	16.8	84.0	35409	3	CEC44C10	Z69787 Caenorhabdi
11	c	16.8	84.0	36727	9	AL512289	AL512289 Human DNA
12	c	16.8	84.0	37154	3	AF025467	AF025467 Caenorhab
13	c	16.8	84.0	71198	2	AC127624	AC127624 Rattus no
14	c	16.8	84.0	76922	5	AC087104	AC087104 Danio rer
15	c	16.8	84.0	95594	2	AC122615	AC122615 Rattus no
16		16.8	84.0	96075	2	AC018294	AC018294 Drosophil
17	c	16.8	84.0	97242	8	AC051630	AC051630 Arabidops
18		16.8	84.0	97310	2	AC103083	AC103083 Rattus no
19	c	16.8	84.0	107658	2	AC119592	AC119592 Rattus no
20		16.8	84.0	108713	2	AC108963	AC108963 Rattus no
21		16.8	84.0	110000	2	CEY111B2_2	Continuation (3 of
22	c	16.8	84.0	115530	2	AC120980	AC120980 Rattus no
23	c	16.8	84.0	115984	9	HS292F10	AL008631 Human DNA
24		16.8	84.0	127695	2	AC111722	AC111722 Rattus no
25		16.8	84.0	139838	2	AL357652	AL357652 Homo sapi
26		16.8	84.0	141908	2	AC123282	AC123282 Rattus no
27		16.8	84.0	142157	2	AC102412	AC102412 Mus muscu
28		16.8	84.0	146056	2	AC128899	AC128899 Rattus no
29		16.8	84.0	152854	2	AC110514	AC110514 Mus muscu
30	c	16.8	84.0	159933	2	AC109935	AC109935 Rattus no
31	c	16.8	84.0	164583	2	AC098073	AC098073 Rattus no
32		16.8	84.0	165454	2	AC113097	AC113097 Mus muscu
33	c	16.8	84.0	169620	2	AC012674	AC012674 Homo sapi
34	c	16.8	84.0	169656	2	AC097292	AC097292 Rattus no
35		16.8	84.0	173548	2	AC094369	AC094369 Rattus no
36		16.8	84.0	173906	2	AC121059	AC121059 Rattus no
37		16.8	84.0	177786	2	AC093609	AC093609 Homo sapi
38		16.8	84.0	179260	3	AC007770	AC007770 Drosophil
39		16.8	84.0	182730	9	AC016735	AC016735 Homo sapi
40	c	16.8	84.0	184258	2	AC117231	AC117231 Mus muscu
41	c	16.8	84.0	184380	10	AC124521	AC124521 Mus muscu
42	c	16.8	84.0	192761	2	AC125874	AC125874 Rattus no
43		16.8	84.0	196785	2	AC120065	AC120065 Rattus no
44	c	16.8	84.0	199289	9	AC012378	AC012378 Homo sapi
45		16.8	84.0	236774	2	AL772275	AL772275 Mus muscu

ALIGNMENTS

RESULT 1
A89795
LOCUS A89795 20 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 17 from Patent WO9832462.
ACCESSION A89795
VERSION A89795.1 GI:6738309
KEYWORDS .
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Lipford,G.B. and Heeg,K.
TITLE PHARMACEUTICAL COMPOSITIONS COMPRISING A POLYNUCLEOTIDE AND OPTIONALLY AN ANTIGEN ESPECIALLY FOR VACCINATION
JOURNAL Patent: WO 9832462-A 17 30-JUL-1998;

```
          LIPFORD GRAYSON B (DE); HEEG KLAUS (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        4 a      6 c      2 g      8 t
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTCCATTTCCCGTAAATCTT 20
              ||||||||||||||||||||
Db          1 GTCCATTTCCCGTAAATCTT 20


RESULT 2
A90882
LOCUS       A90882                   20 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 17 from Patent EP0855184.
ACCESSION   A90882
VERSION     A90882.1  GI:6739307
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Heeg,K.P. and Lipford,G.B.
  TITLE     Pharmaceutical composition comprising a polynucleotide and an
            antigen especially for vaccination
  JOURNAL   Patent: EP 0855184-A 17 29-JUL-1998;
            HEEG KLAUS PROF DR (DE); LIPFORD GRAYSON B DR (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        4 a      6 c      2 g      8 t
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTCCATTTCCCGTAAATCTT 20
              ||||||||||||||||||||
Db          1 GTCCATTTCCCGTAAATCTT 20


RESULT 3
AX455587
LOCUS       AX455587                 20 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 64 from Patent WO0222809.
ACCESSION   AX455587
VERSION     AX455587.1  GI:21714655
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bauer,S., Lipford,G. and Wagner,H.
  TITLE     Process for high throughput screening of cpg-based
            immuno-agonist/antagonist
  JOURNAL   Patent: WO 0222809-A 64 21-MAR-2002;
            Coley Pharmaceutical GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic oligonucleotide"
BASE COUNT        4 a      6 c      2 g      8 t
ORIGIN

  Query Match           100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GTCCATTTCCCGTAAATCTT 20
              ||||||||||||||||||||
Db          1 GTCCATTTCCCGTAAATCTT 20


RESULT 4
AC104070
LOCUS       AC104070               123291 bp    DNA     linear   PRI 29-MAY-2002
DEFINITION  Homo sapiens BAC clone RP11-279K24 from 4, complete sequence.
ACCESSION   AC104070 AC068461
VERSION     AC104070.3  GI:20279508
KEYWORDS    HTG.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 123291)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 123291)
  AUTHORS   Desai,A., Kozlowicz,A. and Boyer,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-279K24
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 123291)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-DEC-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 123291)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-APR-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 123291)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Apr 24, 2002 this sequence version replaced gi:18030153.
            ---------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            ---------- Summary Statistics
                Center project name: H_NH0279K24
                Drafting Center: WIBR
            ----------

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
```

? Nov, 98.

McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-45I20, 2000 bp overlap; the clone sequenced to the right is RP11-173M11. Actual start of this clone is at base position 108871 of RP11-45I20; actual end is at base position 123291 of RP11-279K24.

Unresolved tandem repeats exist between 44681 and 46316.
Polymorphisms exist between AC096659, AC0110771 and AC104070. Data from AC110771 was used to finish AC104070.

The sequence of AC068461 has been incorporated into AC104070.

FEATURES Location/Qualifiers
source 1. .123291
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-279K24"
/clone_lib="RPCI-11"
repeat_region 174. .350
/rpt_family="MER1_type"
repeat_region 1085. .1129
/rpt_family="(CA)n"
repeat_region 2938. .3252
/rpt_family="Alu"
repeat_region 3691. .3908
/rpt_family="L2"
repeat_region 4040. .4334
/rpt_family="Alu"
repeat_region 4345. .4522
/rpt_family="Alu"
repeat_region 4525. .5166
/rpt_family="ERV1"
repeat_region 5269. .5343
/rpt_family="MIR"
repeat_region 5469. .5535
/rpt_family="MIR"
repeat_region 5866. .6417
/rpt_family="ERVL"
repeat_region 7027. .7321
/rpt_family="L2"
repeat_region 7331. .7424
/rpt_family="L2"
repeat_region 7535. .7724
/rpt_family="ERV1"
repeat_region 7739. .8079
/rpt_family="ERV1"
repeat_region 11056. .11367
/rpt_family="Alu"
repeat_region 12482. .12789
/rpt_family="Alu"
repeat_region 12790. .12814
/rpt_family="AT_rich"
repeat_region 13276. .13375
/rpt_family="L2"
repeat_region 13504. .13774
/rpt_family="L1"
repeat_region 13776. .14086
/rpt_family="Alu"
repeat_region 14090. .14277
/rpt_family="L1"
repeat_region 14307. .14427
/rpt_family="L1"
repeat_region 14642. .14709
/rpt_family="U4"
repeat_region 14802. .14883
/rpt_family="CT-rich"
repeat_region 15149. .15220
/rpt_family="ERV1"
repeat_region 15221. .15419
/rpt_family="MaLR"
repeat_region 15420. .16124
/rpt_family="ERV1"
repeat_region 16124. .16356
/rpt_family="MaLR"
repeat_region 16357. .16729
/rpt_family="ERV1"
repeat_region 16730. .17029
/rpt_family="Alu"
repeat_region 17030. .17494
/rpt_family="ERV1"
repeat_region 17517. .17607
/rpt_family="L2"
repeat_region 18972. .19151
/rpt_family="ERV1"
repeat_region 19204. .19453
/rpt_family="ERV1"
repeat_region 19564. .19640
/rpt_family="GA-rich"
repeat_region 20193. .21044
/rpt_family="ERV1"
repeat_region 21501. .21792
/rpt_family="Alu"
repeat_region 23724. .23747
/rpt_family="AT_rich"
repeat_region 24122. .24419
/rpt_family="Alu"
repeat_region 24525. .24920
/rpt_family="MaLR"
repeat_region 25364. .25788
/rpt_family="ERV1"
repeat_region 26725. .27020
/rpt_family="Alu"
repeat_region 27333. .27551
/rpt_family="MIR"
repeat_region 28646. .28951
/rpt_family="Alu"
repeat_region 28952. .29088
/rpt_family="Alu"
repeat_region 29204. .29740
/rpt_family="L1"
repeat_region 29741. .30050
/rpt_family="Alu"
repeat_region 30051. .30244
/rpt_family="L1"
repeat_region 30245. .30462
/rpt_family="Alu"
repeat_region 30463. .30556
/rpt_family="L1"
repeat_region 30733. .30767
/rpt_family="U2"
repeat_region 30768. .30845
/rpt_family="L1"
repeat_region 32081. .32145

Query Match 95.0%; Score 19; DB 9; Length 123291;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TCCATTTCCCGTAAATCTT 20
||||||||||||||||||||
Db 18891 TCCATTTCCCGTAAATCTT 18909

```
RESULT 5
AX040434
LOCUS       AX040434                  22 bp    DNA     linear   PAT 18-NOV-2000
DEFINITION  Sequence 9 from Patent WO0063357.
ACCESSION   AX040434
VERSION     AX040434.1  GI:11230241
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Flier,J.S. and Bjorbaek,C.
  TITLE     Methods and compositions for modulating ciliary neurotrophic facto
            r activity
  JOURNAL   Patent: WO 0063357-A 9 26-OCT-2000;
            Beth Israel Deaconess Medical Center (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer-bind"
BASE COUNT         5 a      8 c      2 g      7 t
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 6;  Length 22;
  Best Local Similarity   94.7%;  Pred. No. 1.5e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TCCATTTCCCGTAAATCTT 20
              ||||||||||||||||||
Db          4 TCCATTTCCCGTAAATCAT 22


RESULT 6
AC094579/c
LOCUS       AC094579                105428 bp    DNA     linear   HTG 11-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-4M8, *** SEQUENCING IN PROGRESS ***,
            56 unordered pieces.
ACCESSION   AC094579
VERSION     AC094579.3  GI:21717874
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 105428)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            [illegible]J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            [illegible]o,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            [illegible]ari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 105428)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 105428)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Jul 10, 2002 this sequence version replaced gi:17941341.
            ---------- Genome Center
               Center: Baylor College of Medicine
               Center code: BCM
               Web site: http://www.hgsc.bcm.tmc.edu/
               Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
               Center project name: GAYF
               Center clone name: CH230-4M8
            ---------- Summary Statistics
               Sequencing vector: Plasmid;
               Chemistry: Dye-terminator Big Dye: 100% of reads
               Assembly program: Phrap; version 0.990329
               Consensus quality: 28157 bases at least Q40
               Consensus quality: 31757 bases at least Q30
               Consensus quality: 34978 bases at least Q20
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 56 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1     1042: contig of 1042 bp in length
            *       1043     1142: gap of unknown length
            *       1143     2198: contig of 1056 bp in length
            *       2199     2298: gap of unknown length
            *       2299     3393: contig of 1095 bp in length
            *       3394     3493: gap of unknown length
            *       3494     4541: contig of 1048 bp in length
            *       4542     4641: gap of unknown length
            *       4642     5660: contig of 1019 bp in length
            *       5661     5760: gap of unknown length
            *       5761     7084: contig of 1324 bp in length
            *       7085     7184: gap of unknown length
            *       7185     8421: contig of 1237 bp in length
            *       8422     8521: gap of unknown length
            *       8522     9745: contig of 1224 bp in length
            *       9746     9845: gap of unknown length
```

```
*     9846    11154: contig of 1309 bp in length
*    11155    11254: gap of unknown length
*    11255    12679: contig of 1425 bp in length
*    12680    12779: gap of unknown length
*    12780    13781: contig of 1002 bp in length
*    13782    13881: gap of unknown length
*    13882    15487: contig of 1606 bp in length
*    15488    15587: gap of unknown length
*    15588    17254: contig of 1667 bp in length
*    17255    17354: gap of unknown length
*    17355    18921: contig of 1567 bp in length
*    18922    19021: gap of unknown length
*    19022    20401: contig of 1380 bp in length
*    20402    20501: gap of unknown length
*    20502    21923: contig of 1422 bp in length
*    21924    22023: gap of unknown length
*    22024    23458: contig of 1435 bp in length
*    23459    23558: gap of unknown length
*    23559    24766: contig of 1208 bp in length
*    24767    24866: gap of unknown length
*    24867    26206: contig of 1340 bp in length
*    26207    26306: gap of unknown length
*    26307    27501: contig of 1195 bp in length
*    27502    27601: gap of unknown length
*    27602    29040: contig of 1439 bp in length
*    29041    29140: gap of unknown length
*    29141    30213: contig of 1073 bp in length
*    30214    30313: gap of unknown length
*    30314    32113: contig of 1800 bp in length
*    32114    32213: gap of unknown length
*    32214    33613: contig of 1400 bp in length
*    33614    33713: gap of unknown length
*    33714    35543: contig of 1830 bp in length
*    35544    35643: gap of unknown length
*    35644    37465: contig of 1822 bp in length
*    37466    37565: gap of unknown length
*    37566    39308: contig of 1743 bp in length
*    39309    39408: gap of unknown length
*    39409    41144: contig of 1736 bp in length
*    41145    41244: gap of unknown length
*    41245    43190: contig of 1946 bp in length
*    43191    43290: gap of unknown length
*    43291    44853: contig of 1563 bp in length
*    44854    44953: gap of unknown length
*    44954    46524: contig of 1571 bp in length
*    46525    46624: gap of unknown length
*    46625    48160: contig of 1536 bp in length
*    48161    48260: gap of unknown length
*    48261    49738: contig of 1478 bp in length
*    49739    49838: gap of unknown length
*    49839    51496: contig of 1658 bp in length
*    51497    51596: gap of unknown length
*    51597    53511: contig of 1915 bp in length
*    53512    53611: gap of unknown length
*    53612    54616: contig of 1005 bp in length
*    54617    54716: gap of unknown length
*    54717    55994: contig of 1278 bp in length
*    55995    56094: gap of unknown length
*    56095    57817: contig of 1723 bp in length
*    57818    57917: gap of unknown length
*    57918    59237: contig of 1320 bp in length
*    59238    59337: gap of unknown length
*    59338    61841: contig of 2504 bp in length
*    61842    61941: gap of unknown length
*    61942    63171: contig of 1230 bp in length
*    63172    63271: gap of unknown length
*    63272    64768: contig of 1497 bp in length
*    64769    64868: gap of unknown length
*    64869    66113: contig of 1245 bp in length
*    66114    66213: gap of unknown length
*    66214    67915: contig of 1702 bp in length
*    67916    68015: gap of unknown length
*    68016    69570: contig of 1555 bp in length
*    69571    69670: gap of unknown length
*    69671    72019: contig of 2349 bp in length
*    72020    72119: gap of unknown length
*    72120    74669: contig of 2550 bp in length
*    74670    74769: gap of unknown length
*    74770    76723: contig of 1954 bp in length
*    76724    76823: gap of unknown length
*    76824    80072: contig of 3249 bp in length
*    80073    80172: gap of unknown length
*    80173    82228: contig of 2056 bp in length
*    82229    82328: gap of unknown length
*    82329    85455: contig of 3127 bp in length
*    85456    85555: gap of unknown length
*    85556    88149: contig of 2594 bp in length
*    88150    88249: gap of unknown length
*    88250    91553: contig of 3304 bp in length
*    91554    91653: gap of unknown length
*    91654    94432: contig of 2779 bp in length

Query Match             87.0%;  Score 17.4;  DB 2;  Length 105428;
Best Local Similarity   94.7%;  Pred. No. 4.2e+02;
Matches   18;  Conservative    0;  Mismatches   1;  Indels   0;  Gaps   0;

Qy          2 TCCATTTCCCGTAAATCTT 20
              ||||||||||||| ||||||
Db      96461 TCCATTTCCCGTTAATCTT 96443


RESULT 7
AC127444/c
LOCUS       AC127444               164383 bp    DNA     linear   HTG 31-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-254F10, *** SEQUENCING IN PROGRESS
            ***, 58 unordered pieces.
ACCESSION   AC127444
VERSION     AC127444.2  GI:21953857
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 164383)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
```

Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
TITLE Weinstock,G. and Gibbs,R.
JOURNAL Direct Submission
Unpublished
REFERENCE 2 (bases 1 to 164383)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 164383)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jul 24, 2002 this sequence version replaced gi:21886827.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GNRV
Center clone name: CH230-254F10
---------- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 110846 bases at least Q40
Consensus quality: 116562 bases at least Q30
Consensus quality: 120375 bases at least Q20
----
* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 58 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1068: contig of 1068 bp in length
1069	1168: gap of unknown length
1169	2250: contig of 1082 bp in length
2251	2350: gap of unknown length
2351	3542: contig of 1192 bp in length
3543	3642: gap of unknown length
3643	4661: contig of 1019 bp in length
4662	4761: gap of unknown length
4762	5958: contig of 1197 bp in length
5959	6058: gap of unknown length
6059	7097: contig of 1039 bp in length
7098	7197: gap of unknown length
7198	8623: contig of 1426 bp in length
8624	8723: gap of unknown length
8724	10322: contig of 1599 bp in length
10323	10422: gap of unknown length
10423	11682: contig of 1260 bp in length
11683	11782: gap of unknown length
11783	13597: contig of 1815 bp in length
13598	13697: gap of unknown length
13698	14894: contig of 1197 bp in length
14895	14994: gap of unknown length
14995	16266: contig of 1272 bp in length
16267	16366: gap of unknown length
16367	17583: contig of 1217 bp in length
17584	17683: gap of unknown length
17684	18852: contig of 1169 bp in length
18853	18952: gap of unknown length
18953	20899: contig of 1947 bp in length
20900	20999: gap of unknown length
21000	22340: contig of 1341 bp in length
22341	22440: gap of unknown length
22441	23863: contig of 1423 bp in length
23864	23963: gap of unknown length
23964	25333: contig of 1370 bp in length
25334	25433: gap of unknown length
25434	27590: contig of 2157 bp in length
27591	27690: gap of unknown length
27691	30206: contig of 2516 bp in length
30207	30306: gap of unknown length
30307	32493: contig of 2187 bp in length
32494	32593: gap of unknown length
32594	34681: contig of 2088 bp in length
34682	34781: gap of unknown length
34782	36994: contig of 2213 bp in length
36995	37094: gap of unknown length
37095	38293: contig of 1199 bp in length
38294	38393: gap of unknown length
38394	40151: contig of 1758 bp in length
40152	40251: gap of unknown length
40252	43157: contig of 2906 bp in length
43158	43257: gap of unknown length
43258	44695: contig of 1438 bp in length
44696	44795: gap of unknown length
44796	47432: contig of 2637 bp in length
47433	47532: gap of unknown length
47533	49807: contig of 2275 bp in length
49808	49907: gap of unknown length
49908	52732: contig of 2825 bp in length
52733	52832: gap of unknown length
52833	54682: contig of 1850 bp in length
54683	54782: gap of unknown length
54783	57132: contig of 2350 bp in length
57133	57232: gap of unknown length
57233	59753: contig of 2521 bp in length
59754	59853: gap of unknown length
59854	63234: contig of 3381 bp in length
63235	63334: gap of unknown length
63335	66253: contig of 2919 bp in length
66254	66353: gap of unknown length
66354	69879: contig of 3526 bp in length
69880	69979: gap of unknown length
69980	72675: contig of 2696 bp in length
72676	72775: gap of unknown length
72776	75756: contig of 2981 bp in length
75757	75856: gap of unknown length
75857	79290: contig of 3434 bp in length
79291	79390: gap of unknown length
79391	83502: contig of 4112 bp in length
83503	83602: gap of unknown length
83603	85785: contig of 2183 bp in length
85786	85885: gap of unknown length
85886	89741: contig of 3856 bp in length
89742	89841: gap of unknown length
89842	93785: contig of 3944 bp in length
93786	93885: gap of unknown length
93886	97426: contig of 3541 bp in length
97427	97526: gap of unknown length
97527	101477: contig of 3951 bp in length
101478	101577: gap of unknown length
101578	104361: contig of 2784 bp in length
104362	104461: gap of unknown length
104462	108173: contig of 3712 bp in length
108174	108273: gap of unknown length
108274	112365: contig of 4092 bp in length
112366	112465 gap of unknown length

```
     *   112466   115356: contig of 2891 bp in length
     *   115357   115456: gap of unknown length
     *   115457   119061: contig of 3605 bp in length
     *   119062   119161: gap of unknown length
     *   119162   122758: contig of 3597 bp in length
     *   122759   122858: gap of unknown length
     *   122859   125799: contig of 2941 bp in length
     *   125800   125899: gap of unknown length
     *   125900   129716: contig of 3817 bp in length
     *   129717   129816: gap of unknown length
     *   129817   136172: contig of 6356 bp in length

  Query Match             87.0%;  Score 17.4;  DB 2;  Length 164383;
  Best Local Similarity   94.7%;  Pred. No. 4.4e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TCCATTTCCCGTAAATCTT 20
              ||||||||||||||||||
Db      63019 TCCATTTCCCGTAAATCGT 63001


RESULT 8
I81948
LOCUS       I81948                    19 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 46 from patent US 5712094.
ACCESSION   I81948
VERSION     I81948.1  GI:3210245
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Seidel,H.Martin., Lamb,I.Peter. and Chan,S.-S.Tian.
  TITLE     Methods for detecting modulators of cytokine action
  JOURNAL   Patent: US 5712094-A 46 27-JAN-1998;
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="unknown"

BASE COUNT        5 a      6 c      2 g      6 t
ORIGIN

  Query Match             85.0%;  Score 17;  DB 6;  Length 19;
  Best Local Similarity  100.0%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 TCCATTTCCCGTAAATC 18
              |||||||||||||||||
Db          3 TCCATTTCCCGTAAATC 19


RESULT 9
AF498313/c
LOCUS       AF498313                2274 bp    DNA     linear   BCT 13-JUN-2002
DEFINITION  Streptococcus infantis strain Atu-4 histidine kinase ComD (comD)
            and response regulator ComE (comE) genes, complete cds.
ACCESSION   AF498313
VERSION     AF498313.1  GI:20269892
KEYWORDS    .
SOURCE      Streptococcus infantis.
  ORGANISM  Streptococcus infantis
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 2274)
  AUTHORS   Ween,O., Teigen,S., Gaustad,P., Kilian,M. and Havarstein,L.S.
  TITLE     Competence without a competence pheromone in a natural isolate of
            Streptococcus infantis
  JOURNAL   J. Bacteriol. 184 (13), 3426-3432 (2002)
  MEDLINE   22053206
    PUBMED  12057935
REFERENCE   2  (bases 1 to 2274)
  AUTHORS   Ween,O., Teigen,S., Gaustad,P., Kilian,M. and Havarstein,L.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-APR-2002) Department of Chemistry and Biotechnology,
            Agricultural University of Norway, P.O. Box 5040, Aas, Norway
FEATURES             Location/Qualifiers
     source          1. .2274
                     /organism="Streptococcus infantis"
                     /strain="Atu-4"
                     /db_xref="taxon:68892"
     gene            133. .1452
                     /gene="comD"
     CDS             133. .1452
                     /gene="comD"
                     /codon_start=1
                     /transl_table=11
                     /product="histidine kinase ComD"
                     /protein_id="AAM18090.1"
                     /db_xref="GI:20269893"
                     /translation="MDLLKILIVILHILILTNSYELICKATKKERYLFISWVFILQSI
                     VEILFYLTSLHGLGMEKISFTCIIFAYLIGFKKYNKFKAIFISLLLSLLYQSTHTFIA
                     VIISSLTGDSFLIKYEDLFFVVVLMLTYFVIKKIIFYFHLEINYFDKDYLYPFFKKVI
                     FAFFALHILLFISDMVNIHEHFYSFSSILATIVFICLLLIFFAMNAKKVQIEKEIALK
                     QKKFEQEHLQTYTDEIVGLYNEIRGFRHDYAGMLVSMQMAIDSKDLQEIDRIYNEVLV
                     KANQKLRSDKYTYFDLNNIEDTALRSLIAQSIVYARNKDVEFTLEVKDIITKLSIDLL
                     DLVRIMSVLLNNAVEGAAESYSKQMEVAVIKMDLEIVIVIQNSCKYTMMPSEDLFELG
                     FSTKGRNRGIGLNNVKEILDQYENIILETEMEDSKFRQIIRFKREFE"
     gene            1453. .2205
                     /gene="comE"
     CDS             1453. .2205
                     /gene="comE"
                     /codon_start=1
                     /transl_table=11
                     /product="response regulator ComE"
                     /protein_id="AAM18091.1"
                     /db_xref="GI:20269894"
                     /translation="MKVLILEDVIQHQVRLERTLDEISKELNIPISYKTTGKEREFKE
                     YIENDEVNQLYFLDIDIHGVEKKGFEVAQFIRRHNPYAIIVFITSRSEFATLTYKYQV
                     SALDFVDKDINDELFKKRIEQSILYTKSMLLENKDVVDYFDYNYKGNDLKIPYHDILY
                     IETTGVSHKLRIIGKNFAKEFYGTMADIQEKDKDTQRFYSAHKSFLVNVGNIREIDRK
                     NLEVVFYEDHRCPITRLKVRKLRDILEKKSKK"

BASE COUNT      838 a    259 c    381 g    796 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 1;  Length 2274;
  Best Local Similarity   90.0%;  Pred. No. 5.4e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GTCCATTTCCCGTAAATCTT 20
              ||| ||||||  |||||||||
Db        936 GTCAATTTCCTGTAAATCTT 917


RESULT 10
CEC44C10/c
LOCUS       CEC44C10                35409 bp    DNA     linear   INV 13-AUG-2002
DEFINITION  Caenorhabditis elegans cosmid C44C10, complete sequence.
ACCESSION   Z69787
VERSION     Z69787.1  GI:3869219
KEYWORDS    HTG; Chromosome segregation protein like; Collagen; Cytochrome
            P450; DNA repair protein RECN like; Protein kinase.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1
  AUTHORS   none.
  TITLE     Genome sequence of the nematode C. elegans: a platform for
            investigating biology. The C. elegans Sequencing Consortium
  JOURNAL   Science 282 (5396), 2012-2018 (1998)
  MEDLINE   99069613
  REMARK    The C.elegans Sequencing Consortium.
REFERENCE   2  (bases 1 to 35409)
  AUTHORS   Cottage,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-1996) Nematode Sequencing Project, Sanger
            Institute, Hinxton, Cambridge CB10 1SA, England and Department of
```

```
COMMENT     Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
            jes@sanger.ac.uk or rw@nematode.wustl.edu
            On Nov 13, 1998 this sequence version replaced gi:1212826.
            Coding sequences below are predicted from computer analysis, using
            predictions from Genefinder (P. Green, U. Washington), and other
            available information.
            Current sequence finishing criteria for the C. elegans genome
            sequencing consortium are that all bases are either sequenced
            unambiguously on both strands, or on a single strand with both a
            dye primer and dye terminator reaction, from distinct subclones.
            Exceptions are indicated by an explicit note.
            IMPORTANT: This sequence is not the entire insert of clone C44C10.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we arrange for a small overlap between
            neighbouring submissions.
            The true left end of clone C44C10 is at 28483 in
            sequence Z66565.
            The true right end of clone C44C10 is at 35409 in this sequence.
            The true left end of clone D1053 is at 33388 in this sequence. The
            true right end of clone T04F8 is at 115 in this sequence. The start
            of this sequence (1. .115) overlaps with the end of sequence Z66565.
            The end of this sequence (35347. .35409) overlaps with the start of
            sequence Z66560.
            [981006 dl] : Cosmid flipped
            For a graphical representation of this sequence and its analysis
            see:- http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?
            name=C44C10
            IMPORTANT:  This sequence is NOT necessarily the entire insert of
            the specified clone.  It may be shorter because we only sequence
            overlapping sections once, or longer because we arrange for a small
            overlap between neighbouring submissions.
FEATURES            Location/Qualifiers
     source         1. .35409
                    /organism="Caenorhabditis elegans"
                    /db_xref="taxon:6239"
                    /chromosome="X"
     gene           /clone="C44C10"
     CDS            join(1823. .1965,2023. .2137,2272. .2467,2515. .2741)
                    /gene="C44C10.8"
                    join(1823. .1965,2023. .2137,2272. .2467,2515. .2741)
                    /gene="C44C10.8"
                    /note="predicted using Genefinder
                    contains similarity to Pfam domain: PF00010
                    (Helix-loop-helix DNA-binding domain), Score=40.1,
                    E-value=1.6e-08, N=1"
                    /codon_start=1
                    /protein_id="CAA93634.1"
                    /db_xref="GI:3874976"
                    /db_xref="SPTREMBL:Q18612"
                    /translation="MVKSTTAGNNAVSSLESTDSKKSRKEKSREKEHRRAQCINSAFE
                    ILQQHIPYLKSEERKSLPKIKTLRLAMQYIDHLKKLLGGNEMLETDCNESRPLTHSDF
                    RVNISNEIRIRNSYRERAHNQELDEATVKRILAREDQRRRCSSVPEGIQRYTPYQSRQ
                    FGPISNNVCNFRHVDANQYNTNFMEYQTQMVQVANFPNQFSHEYYNYGTFPVIDPTSL
     gene           ENSNPILQ"
                    complement(join(4096. .4284,4333. .4587,4631. .4818,
     CDS            4871. .5147,5249. .5365))
                    /gene="C44C10.7"
                    complement(join(4096. .4284,4333. .4587,4631. .4818,
                    4871. .5147,5249. .5365))
                    /gene="C44C10.7"
                    /note="predicted using Genefinder
                    Similarity to protein kinases, contains similarity to Pfam
                    domain: PF00069 (Eukaryotic protein kinase domain),
                    Score=43.0, E-value=2.4e-11, N=1"
                    /codon_start=1
                    /protein_id="CAA93637.1"
                    /db_xref="GI:3874979"
                    /db_xref="SPTREMBL:Q18615"
                    /translation="MSPNSLFQSLQLFHKLSISLILFLHHFSCFLTDLQSSKMTFPIV
                    GDNVRNYEIISVLGGGAFGTVYSCVDVDDRLNQLAIKAMDLSTVTRENSYKLELMVLQ
                    RVETLSEVEQTRFSKLIGNFIQDSSLGFFVMHKEGECVDEVWMRNKSGRFSASNVLKI
                    VHCMAHGLRSLHKIGFIHRDCHAGNILFASRLTPTAPVKIVDFGIGRRFANRRGHPIP
                    SPSRDIDFLGCEHCSVNVHVGGIPGPKDDFLSMVLLAIRMSGIHAISYENSETNLCQK
                    KSFESNPAQFLMRCPWLVPVSVAIIKNDPNRFNYDCVIHAIENASLFNPSEVIGHCYE
     gene           NGLLRVF"
                    join(7408. .7548,7641. .7883)
     CDS            /gene="C44C10.10"
                    join(7408. .7548,7641. .7883)
                    /gene="C44C10.10"
                    /note="Similarity to Bacilus subtilis DNA repair protein
                    RECN (SW:RECN_BACSU)"
                    /codon_start=1
                    /protein_id="CAA93640.1"
                    /db_xref="GI:3874982"
                    /db_xref="SPTREMBL:Q18618"
                    /translation="MDREITTRKRCCKTTFTALSITHHFIYLFLPLSMSASGGFIIKL
                    IDSMTQLLSISVKNFMCLEDVKYESFGKITTITGPNGCGKSALVRAIGILTGDEVTES
     gene           SNQQYAVRFGKYINFWYLWVSFPPK"
                    join(8203. .8384,8502. .8580,8700. .8949,9003. .9061,
     CDS            9106. .9306,9404. .9550)
                    /gene="C44C10.5"
                    join(8203. .8384,8502. .8580,8700. .8949,9003. .9061,
                    9106. .9306,9404. .9550)
                    /gene="C44C10.5"
                    /note="Similarity to Yeast chromosome segregation protein
                    SMC-1 (SW:SMC1_YEAST)"
                    /codon_start=1
                    /protein_id="CAA93639.1"
                    /db_xref="GI:3874981"
                    /db_xref="SPTREMBL:Q18617"
                    /translation="MVDNLHRTREEYLRELDDITLRVHGRKLSVYSSSSKVTQDNAED
                    FGLFLEAIFDSIEHKPKLFLNDLERKSLEKKLTKEQVKLTNIKASKQSQKVTKSDDPK
                    AIAKQAPPAMFKVQGCSRLDGDSLSQENAEKPAIVIELENRLVQQRAQRRKIIANVTL
                    QDLDIPAKEAGTFSYDVKPDYSSLPDELKKESNDAETVKMHLRTLEETIQATRNEWVA
     gene           VNGAEYPDMSNILIRFHETDSFINIAKLRTTELEAEIEQVKASRRTHFNQRLAAVSSS
                    VSRIYREISEDPDSTVSLASSNEDEPYEC"
                    join(10560. .10757,10802. .11290,11372. .11609,11658. .11716,
     CDS            11765. .11968,12057. .12377,12582. .12686)
                    /gene="C44C10.4"
                    join(10560. .10757,10802. .11290,11372. .11609,11658. .11716,
                    11765. .11968,12057. .12377,12582. .12686)
                    /gene="C44C10.4"
                    /note="Similarity to Yeast chromosome segregation protein
                    SMC-1 (SW:SMC1_YEAST)
                    cDNA EST EMBL:BJ138908 comes from this gene"
                    /codon_start=1
                    /protein_id="CAA93638.2"
                    /db_xref="GI:22265835"
                    /db_xref="SPTREMBL:Q18616"
                    /translation="MTYPQQRLLTVDILNFKCYQHHKSFGPFDHVTTFIGMNGCGKSI
                    MIEAIAFVFGEDLPDPKNPHNEHLFYDYNIRCVVSLKFEMPDGSTKRFTRIGLKTTYT
                    YMIFNKYVTEEEYLLELGKMNFRFAGQNISVYNKTTLKVTRKSAEELGAFFDIVSNSI
                    EFKPEYDELKAKLNFAESTVLRWAKMIDSIELEKEKFITLTLYKLYLCDKEIQETKEK
     gene           FVAEQLKVNAKNAKMPLRISRYSEAIPKQTPSPNFKNNAIVVRPKTHNHYDNLITVDK
                    LNSYLNEKVTERRELLQFAVNHPNMNIRFSEEGPNIYAREIDYTELPEYYKKESDDII
     CDS            ATSYKRSMLEKEIKKIRQCWLDENNIKDDANIDNINMRIKEATTKHTLALNIESEARL
                    NWEKLRVMRSARFNNFMAPLQTSLSRIYAEIAQDEDSLAFLRPLDEEEPYLAIDAFVK
                    ERHSEHISTRTFHGGSRMIVDLALIFAIHEWKPSPLVVIDDIDRNLYHEASKKIEPYL
     gene           SRQKQTQFIVTTKEDSYFTEVTKTITLEN"
     CDS            join(15478. .15499,15592. .15765,15953. .16014,16061. .16192)
                    /gene="C44C10.6"
                    join(15478. .15499,15592. .15765,15953. .16014,16061. .16192)
                    /gene="C44C10.6"
                    /codon_start=1
                    /protein_id="CAA93641.1"
                    /db_xref="GI:3874983"
                    /db_xref="SPTREMBL:Q18619"
                    /translation="MRHDVNPFLSRNFNQSTMKLEKTKQSDDYYKKWTREQLCEHAEP
                    LRNLKNPFADMGTFLFSTSNFVLEFPHSVSMFLTKVFRRVAAIESKKCIVGEPSSKCF
     gene           FKSTDEHAFTGNRFTRNRTPYFGILEP"
                    join(19631. .19848,20165. .20420,20559. .20954,20999. .21256,
     CDS            21300. .21386)
                    /gene="C44C10.3"
                    join(19631. .19848,20165. .20420,20559   20954,20999   21256,
                    21300. .21386)
                    /gene="C44C10.3"
```

```
                     /note="predicted using Genefinder
                     contains similarity to Pfam domain: PF00083 (Sugar (and
                     other) transporter), Score=-98.8, E-value=0.00078, N=1"
                     /codon_start=1
                     /protein_id="CAA93635.1"
                     /db_xref="GI:3874977"
                     /db_xref="SPTREMBL:Q18613"
                     /translation="MFYDVNRFHLCVFVTWQLVNFAVGQYIFNIFSIHVPKWKCGDGP
                     VTKDCKVYQSCPKDNLTFVDPIFTSVAMEFCSGLSIAGVLVCICAWILESILSHQRMV
                     IRAFFNWGWTRVFITVVCYLTQEWRLASYVCAACLVPALLMLIFLIPESAIWLHSKGE
                     TAKMIESEKFMAKIAGVSYVPIEHKLIHSKTLLETLNTKGLFKKLFVLWIMWFFVAVC
                     GFANDLNSNSLAGNLYLNQALFGILLVLSKIVLLFVDTKFENFKRRTLHQGSQSGMII
                     SFLILAIFLWIDYHGTGFLIVYLFGTMFIEYTWDAVYLCAIESMETSSRASAVGSCSL
                     MARIGSLLAPFLTYANTWWPPAVYFTIIVLGTVNLIISFMFLQETKNVNLDEVQVNDD
                     EAQEEKIAMLQT"
gene                 complement(join(21983. .22039,22085. .22371,22741. .22945,
                     22993. .23172,23659. .23847,23892. .24010,25399. .26203,
                     26251. .26301,26703. .26846,27118. .27279))
                     /gene="C44C10.2"
CDS                  complement(join(21983. .22039,22085. .22371,22741. .22945,
                     22993. .23172,23659. .23847,23892. .24010,25399. .26203,
                     26251. .26301,26703. .26846,27118. .27279))
                     /gene="C44C10.2"
                     /note="Similarity to Drosophila cytochrome P450 (PIR Acc.
                     No. S34291), contains similarity to Pfam domain: PF00067
                     (Cytochrome P450), Score=155.8, E-value=9.4e-44, N=2
                     cDNA EST EMBL:AU211318 comes from this gene
                     cDNA EST EMBL:AU199850 comes from this gene"
                     /codon_start=1
                     /protein_id="CAA93636.1"
                     /db_xref="GI:3874978"
                     /db_xref="SPTREMBL:Q27480"

Query Match              84.0%;  Score 16.8;  DB 3;  Length 35409;
Best Local Similarity    90.0%;  Pred. No. 7.6e+02;
Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GTCCATTTCCCGTAAATCTT 20
              ||||| ||| ||||||||||
Db      10892 GTCCAATTCTCGTAAATCTT 10873


RESULT 11
AL512289/c
LOCUS       AL512289              36727 bp    DNA     linear   PRI 10-APR-2002
DEFINITION  Human DNA sequence from clone RP11-256L9 on chromosome 6, complete
            sequence.
ACCESSION   AL512289
VERSION     AL512289.18  GI:20145629
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Lloyd,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 12, 2002 this sequence version replaced gi:16944069.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-256L9 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6.
FEATURES                Location/Qualifiers
     source             1. .36727
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="6"
                        /clone="RP11-256L9"
                        /clone_lib="RPCI-11.1"
BASE COUNT     11166 a    6909 c    7196 g   11456 t
ORIGIN

  Query Match              84.0%;  Score 16.8;  DB 9;  Length 36727;
  Best Local Similarity    90.0%;  Pred. No. 7.6e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GTCCATTTCCCGTAAATCTT 20
              | |||||||||| ||||||||
Db      18407 GACCATTTCCCCTAAATCTT 18388


RESULT 12
AF025467/c
LOCUS       AF025467              37154 bp    DNA     linear   INV 31-MAY-2002
DEFINITION  Caenorhabditis elegans cosmid R148, complete sequence.
ACCESSION   AF025467
VERSION     AF025467.1  GI:2429504
KEYWORDS    HTG.
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 37154)
  AUTHORS   Waterston,R.
  TITLE     Genome sequence of the nematode C. elegans: a platform for
            investigating biology. The C. elegans Sequencing Consortium
  JOURNAL   Science 282 (5396), 2012-2018 (1998)
  MEDLINE   99069613
   PUBMED   9851916
REFERENCE   2  (bases 1 to 37154)
  AUTHORS   Le,T.T., Kemp,K. and Scheet,P.
  TITLE     The sequence of C. elegans cosmid R148
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 37154)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-1997) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE   4  (bases 1 to 37154)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) Department of Genetics, Washington
            University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
            Louis, MO 63110, USA
REFERENCE   5  (bases 1 to 37154)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-2002) Department of Genetics, Washington
            University, Genome Sequencing Center, 4444 Forest Park Avenue, St
            Louis, MO 63110, USA
REFERENCE   6  (bases 1 to 37154)
```

```
AUTHORS   Waterston,R.
TITLE     Direct Submission
JOURNAL   Submitted (31-MAY-2002) Department of Genetics, Washington
          University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
          Louis, MO 63110, USA
COMMENT   Submitted by:
               Genome Sequencing Center
               Department of Genetics, Washington University
               St. Louis , MO  63110, USA, and
               Sanger Centre, Hinxton Hall
               Cambridge CB10 IRQ, England
               email: rw@nematode.wustl.edu and jes@sanger.ac.uk

          NOTICE:  This sequence may not be the entire insert of this clone.
          It may be shorter because we only sequence overlapping sections
          once, or longer because we provide a small overlap between
          neighboring submissions.

          This sequence was finished as follows unless otherwise noted: all
          regions were double stranded, sequenced with an alternate chemistry
          or covered by high quality data (i.e., phred quality >= 30); an
          attempt was made to resolve all sequencing problems, such as
          compressions and repeats; all regions were covered by sequence from
          more than one m13 subclone.

          For a graphical representation of this cosmid sequence and its
          analysis see:
          http://www.wormbase.org/db/seq/sequence?name=R148;class=Sequence

                        NEIGHBORING COSMID INFORMATION

          The 5' cosmid is C44B11, 900 bp overlap; the 3' cosmid is H09G03,
          2400 bp overlap. Actual start of this cosmid is at base position
          895 of R148; actual end is at 37154 of R148.

          NOTES:

          Coding seqences below are the result of integration and manual
          review of the following data : computer analysis using the program
          Genefinder (P. Green and L. Hillier, personal communication), the
          large scale EST projects of Yuji Kohara
          (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C.
          elegans ORFeome cloning project (http://worfdb.dfci.harvard.edu/),
          similarity to other proteins from BlastX analyses
          (http://blast.wustl.edu/), sequence conservation with C. briggsae
          using Jim Kent's WABA alignment program (Genome Research
          10:1115-1125, 2000), individual C. elegans researchers. tRNAs
          and personal communications with C. elegans GenBank submissions,
          are predicted using the program tRNAscan-SE (Lowe, T.M. and
          Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES
     source          Location/Qualifiers
                     1. .37154
                     /organism="Caenorhabditis elegans"
                     /strain="Bristol N2"
                     /db_xref="taxon:6239"
                     /chromosome="III"
     gene            /clone="R148"
                     95. .4939
                     /gene="R148.5"
                     /note="for a graphical representation of this gene see:
                     http://www.wormbase.org/db/seq/sequence?name=R148.5;class=
     CDS             Sequence"
                     join(95. .435,939. .1303,2277. .2489,3336. .4014,4475. .4559,
                     4844. .4939)
                     /gene="R148.5"
                     /note="contains similarity to drosophila DNA-binding
                     protein K10 (NID:g8148); coded for by the following C.
                     elegans cDNAs: yk211a11.3, yk466a5.3, yk172f2.3,
                     yk185e12.3, yk64h11.3, yk64h11.5, yk79b6.3, yk79b6.5,
                     yk113f3.3, yk113f3.5, yk118e11.3, yk118e11.5, yk172f2.5,
                     yk185e12.5, yk186h12.5, CEEST86F, yk211a11.5, yk461h3.5,
                     yk466a5.5"
                     /codon_start=1
                     /product="Hypothetical protein R148.5"
                     /protein_id="AAB71039.1"
                     /db_xref="GI:2429509"
                     /translation="MVRDRKKLQETKMISVKFSIDWSNKNCKNLKKLFLFKKFFKFFS
                     ARTFLFYFTLLFPKNSNTFQMPLPLCRCRVLYIGSSVPTITKDGLQGIQQPLKERYPI
                     QESPDTRGIDSWLSVWSNGVLLEYIEGEKKTETVFFPIITLHYCAAVRYVNVEGFKVG
                     GGGEQFMPLDSPFANIPNSPHPPIFAAIFRRTTGVKVLECHGFICTNVKAANALVRCM
                     VHAYTETMHLRMDERIPGLKAIKEGSTGDRSPNSPSSEIEENSFEEDVVPEEKNMKNW
                     QERRQQDGEYDSVSISSSLVNKNKKAASEQGGIRGALVPFDEPLRRAGSDNDLSYRGP
                     PPHHPYGPPHPMHHPYAMMPPPFGFFPPPPRGHFPPPPPPHFMGRGMPPPFMPPPPHFG
                     MGPPRGFMPPPHPMMFRGGPYPPFFPPPPPHFMRPSSPPTDGGPIITGPESVYGTMRR
                     GAYEEPAYVGATGGFPEASYQPANYGGDGYETYYDTFKKRGTLKKGESASMASTRADS
     gene            AWDQYEAGIYRKPHINEKAFSGTLRSMAAKETSNATLNRSEHAEVASQTVETEITRPD
                     SPPVDYEAFEKMKLKENSRRAAQPTN"
                     5716. .6352
                     /gene="R148.4"
                     /note="for a graphical representation of this gene see:
                     http://www.wormbase.org/db/seq/sequence?name=R148.4;class=
     CDS             Sequence"
                     join(5716. .5802,5915. .6154,6236. .6352)
                     /gene="R148.4"
                     /codon_start=1
                     /product="Hypothetical protein R148.4"
                     /protein_id="AAB71035.1"
                     /db_xref="GI:2429505"
                     /translation="MSKEANFSFYGALPSFKPTRHCCGIVENSVLYLLLTALTTLQLI
     gene            VFIVIFMHPELRNMLNLNPNLKQNQDILLLILTAMSSIWLLTVLFSFVGFMKKITYSH
                     LPLIATERDLQNSAICRQIRCLRIFDATVPRGIQISSGKPAKNVV"
                     9383. .14111
                     /gene="R148.3"
                     /note="for a graphical representation of this gene see:
                     http://www.wormbase.org/db/seq/sequence?name=R148.3;class=
     CDS             Sequence"
                     join(9383. .9415,9572. .9662,9706. .9842,9887. .10744,
                     10855. .11330,11689. .11841,11895. .13234,13852. .14111)
                     /gene="R148.3"
                     /note="contains weak similarity to human mitotic growth
                     and transcription activator BRG1 protein (GB:S66910);
                     coded for by the following C. elegans cDNAs: yk507a8.5,
                     yk151d12.5, yk168f9.5, yk598c4.5"
                     /codon_start=1
                     /product="Hypothetical protein R148.3"
                     /protein_id="AAB71038.2"
                     /db_xref="GI:15011786"
                     /translation="MEPPKLLGVVLTTALLFHIALTAKQAVVSRKSKLCGNVECEEVL
                     FKARVKRVMGINNDPSFLNLAENAIISVVAVKFSDRPDIMEGKLNGEGSNGFFYAGAI
                     DIMPFAEFLRTAITDKKEMLMISQNPADIGDKRIVGFLHSETDLVRDFNAKNAQAAAE
                     NGLPAPEPIPIPEPAHTGHGHSHGGHGHSHGGHGHSHGHPAPTPAPTPAPKPIEEAPK
                     PAEEPVKPADLEMDAEIERLIKEEEEERMKNVAEQPPLVLQSLPETPEAQKLPEEPLEP
                     VTPEPIPEAPVPIAQYIQEPTTTPAPLEAPAPVENAPNLPTPTFSVDDLMAQAARDAL
                     NSNSIVPDAPKVEEPEPAPTPAPAPAEDPVVTPAPEELTTPPPPAPPPPAPAVPEEAQ
                     APTPAPEEAPEAQIPAETAQIPPENAVPEAPEEPAPTTTPPPISAGEAEFLKWEEARQ
                     KREAVAAAPYEDAQAHQFVASPPPHVSSGLIELIRSSLGLGNITDSSVLLVVNLTFVA
                     ASFLVYIIIRSLSSGGESDILDRQAHFDLMNKHKQLQLQLNAKITEIQNLQANGAAAQ
                     RPEQPNYEPELLALRQQLAHLQSELQTAHAQRQEFEIHYHQIRELAEQKDHQIAQLTE
                     QANDYQERLEKSQSELQELARHVDSARAELMDSQTSDYGRQQELMEKQAELEQIYVKI
                     DNLTAENRDLTDKIAQLQVSDSLELEELEELRNKNLELQEQVKQLKEELEKEKSVAGS
                     GNSGGWSDIDDNEVVEAVQPIETTPISSENSENLKKFEAEIAELKASLEKTEKELTKY
                     RNLYNEKLTAISDKENRSQIEMEHKLKNAESDLQEAKRRAEEIGAEKKELSDRLNEML
                     KNLNQSMQKSAENDKMLTQLREEAFAKEKQLLEHDSVLRKKDDKIRELDAEFKRVKME
                     YVKLETKSFHDVRKLNMDVEELKTQLSMAGGVTSSRSIPDRLVSPLQEPIWDEPEPAL
                     QRVVLPLDGYSSMRRRSTRRSQLPFGAESPSDQEKKQAAPSHRRRSRSHGRQPPSFHD
     gene            PYSMNPYLPGSSDSSFMSTSQPQFLPLNRQHSRSGHLMSNSANGPYSSGGSNGGRSPP
                     PEMPLLSAIPPPGARKPLGKRPGTGEF"
                     17695. .19004
                     /gene="R148.2"
                     /note="for a graphical representation of this gene see:
                     http://www.wormbase.org/db/seq/sequence?name=R148.2;class=
     CDS             Sequence"
                     join(17695. .17729,18063. .18141,18456. .18581,18849. .19004)
                     /gene="R148.2"
                     /note="coded for by the following C  elegans cDNAs
                     CEESX94F"
```

```
                     /codon_start=1
                     /product="Hypothetical protein R148.2"
                     /protein_id="AAB71037.1"
                     /db_xref="GI:2429507"
                     /translation="MESKLEQNVTAMMSSTGASGITVADSEGVRLHSAGKITDMANVG
                     SLMIADAKNMFPKGTRNKKNSYPIVTLHSSDGSKTTKFLFENLKILNVLFKNQQISYY
                     LLQKPEISGFSYKKSVFFSNFLSENLKIF"
                     20702. .25373
     gene            /gene="R148.1"
                     /note="for a graphical representation of this gene see:
                     http://www.wormbase.org/db/seq/sequence?name=R148.1;class=
                     Sequence"
     CDS             join(20702. .20760,21090. .21236,21298. .21405,21460. .21567,
                     22425. .22608,23183. .23258,23309. .23600,24153. .24255,
                     24309. .24353,25110. .25373)
                     /gene="R148.1"

  Query Match             84.0%;  Score 16.8;  DB 3;  Length 37154;
  Best Local Similarity   90.0%;  Pred. No. 7.6e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GTCCATTTCCCGTAAATCTT 20
              || ||||||||||||||| ||
Db      13601 GTGCATTTCCCGTAAATTTT 13582


RESULT 13
AC127624/c
LOCUS       AC127624               71198 bp    DNA     linear   HTG 18-JUL-2002
DEFINITION  Rattus norvegicus clone CH230-263L13, *** SEQUENCING IN PROGRESS
            ***, 41 unordered pieces.
ACCESSION   AC127624
VERSION     AC127624.1  GI:21902925
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 71198)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
            Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
            Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
            Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
            Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
            Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
            Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
            Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
            Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
            Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
            Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
            Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
            Sodergren,E , Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
            Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
            Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
            Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
            Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
            Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
            Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
    TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 71198)
  AUTHORS   Worley,K.C.
    TITLE   Direct Submission
  JOURNAL   Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     ---------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
            Center project name: GXGQ
            Center clone name: CH230-263L13
            ---------- Summary Statistics
            Sequencing vector: Plasmid;
            Chemistry: Dye-terminator Big Dye: 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 25073 bases at least Q40
            Consensus quality: 29679 bases at least Q30
            Consensus quality: 32543 bases at least Q20
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 41 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1023: contig of 1023 bp in length
            *     1024     1123: gap of unknown length
            *     1124     2242: contig of 1119 bp in length
            *     2243     2342: gap of unknown length
            *     2343     3542: contig of 1200 bp in length
            *     3543     3642: gap of unknown length
            *     3643     4763: contig of 1121 bp in length
            *     4764     4863: gap of unknown length
            *     4864     6234: contig of 1371 bp in length
            *     6235     6334: gap of unknown length
            *     6335     7973: contig of 1639 bp in length
            *     7974     8073: gap of unknown length
            *     8074     9541: contig of 1468 bp in length
            *     9542     9641: gap of unknown length
            *     9642    11151: contig of 1510 bp in length
            *    11152    11251: gap of unknown length
            *    11252    12418: contig of 1167 bp in length
            *    12419    12518: gap of unknown length
            *    12519    13596: contig of 1078 bp in length
            *    13597    13696: gap of unknown length
            *    13697    15374: contig of 1678 bp in length
            *    15375    15474: gap of unknown length
            *    15475    16806: contig of 1332 bp in length
            *    16807    16906: gap of unknown length
            *    16907    18166: contig of 1260 bp in length
            *    18167    18266: gap of unknown length
            *    18267    19997: contig of 1731 bp in length
            *    19998    20097: gap of unknown length
            *    20098    21191: contig of 1094 bp in length
            *    21192    21291: gap of unknown length
            *    21292    22497: contig of 1206 bp in length
            *    22498    22597: gap of unknown length
```

```
     22598    24237: contig of 1640 bp in length
     24238    24337: gap of unknown length
     24338    25473: contig of 1136 bp in length
     25474    25573: gap of unknown length
     25574    26800: contig of 1227 bp in length
     26801    26900: gap of unknown length
     26901    28144: contig of 1244 bp in length
     28145    28244: gap of unknown length
     28245    29916: contig of 1672 bp in length
     29917    30016: gap of unknown length
     30017    32125: contig of 2109 bp in length
     32126    32225: gap of unknown length
     32226    34042: contig of 1817 bp in length
     34043    34142: gap of unknown length
     34143    35903: contig of 1761 bp in length
     35904    36003: gap of unknown length
     36004    37019: contig of 1016 bp in length
     37020    37119: gap of unknown length
     37120    38950: contig of 1831 bp in length
     38951    39050: gap of unknown length
     39051    40221: contig of 1171 bp in length
     40222    40321: gap of unknown length
     40322    41557: contig of 1236 bp in length
     41558    41657: gap of unknown length
     41658    42752: contig of 1095 bp in length
     42753    42852: gap of unknown length
     42853    44408: contig of 1556 bp in length
     44409    44508: gap of unknown length
     44509    46146: contig of 1638 bp in length
     46147    46246: gap of unknown length
     46247    47568: contig of 1322 bp in length
     47569    47668: gap of unknown length
     47669    49272: contig of 1604 bp in length
     49273    49372: gap of unknown length
     49373    51788: contig of 2416 bp in length
     51789    51888: gap of unknown length
     51889    54560: contig of 2672 bp in length
     54561    54660: gap of unknown length
     54661    56172: contig of 1512 bp in length
     56173    56272: gap of unknown length
     56273    57893: contig of 1621 bp in length
     57894    57993: gap of unknown length
     57994    60508: contig of 2515 bp in length
     60509    60608: gap of unknown length
     60609    63453: contig of 2845 bp in length
     63454    63553: gap of unknown length
     63554    66714: contig of 3161 bp in length
     66715    66814: gap of unknown length
     66815    71198: contig of 4384 bp in length.
FEATURES             Location/Qualifiers
     source          1. .71198
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-263L13"
BASE COUNT    19367 a  14382 c  13896 g  18404 t   5149 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 2;  Length 71198;
  Best Local Similarity   90.0%;  Pred. No. 8.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GTCCATTTCCCGTAAATCTT 20
              |||||||||| ||| |||||
Db      11923 GTCCATTTCCTGTAGATCTT 11904


RESULT 14
AC087104/c
LOCUS       AC087104               76922 bp    DNA     linear   VRT 04-JUL-2001
DEFINITION  Danio rerio clone 20A7, complete sequence.
ACCESSION   AC087104
VERSION     AC087104.2  GI:14595788
```

```
SOURCE       Danio rerio.
  ORGANISM   Danio rerio
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
             Cypriniformes; Cyprinidae; Danio.
REFERENCE    1  (bases 1 to 76922)
  AUTHORS    Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
             Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L.,
             Granite,S., Guan,X., Gupta,J., Ho,S.-L., Idol,J.R., Karlins,E.,
             Lee-Lin,S.-Q., Legaspi,R., Lim,M., Maduro,Q.L., Maduro,V.B.,
             Masiello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
             Pearson,R., Prasad,A., Shevchenko,Y., Snyder,B., Stantripop,S.,
             Thomas,J.W., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L.,
             Walker,M.A., Wetherby,K.D., Zhang,L.-H. and Green,E.D.
  TITLE      NISC Comparative Sequencing Initiative
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 76922)
  AUTHORS    Green,E.D.
  TITLE      Direct Submission
  JOURNAL    Submitted (07-DEC-2000) NIH Intramural Sequencing Center, 8717
             Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE    3  (bases 1 to 76922)
  AUTHORS    Green,E.D.
  TITLE      Direct Submission
  JOURNAL    Submitted (04-JUL-2001) NIH Intramural Sequencing Center, 8717
             Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT      On Jul 4, 2001 this sequence version replaced gi:11596995.
             -------------- Genome Center
                  Center: NIH Intramural Sequencing Center
                  Center code: NISC
                  Web site: http://www.nisc.nih.gov
                  Contact: nisc_mouse@nhgri.nih.gov
             -------------- Project Information
                  Center project name: aqb
                  Center clone name: 020A07

             This sequence was finished as follows unless otherwise noted:
             all regions were double-stranded, sequenced with an
             alternate chemistry, or covered by high quality data
             (i.e., phred quality >= 30); an attempt was made to resolve
             all sequencing problems, such as compressions and repeats;
             all regions were covered by at least one plasmid subclone
             or more than one M13 subclone; and the assembly was confirmed
             by restriction digest.

             CLONE LENGTH: This sequence represents the entire insert of
             this clone unless otherwise noted. If there are overlapping
             clones, the overlaps are noted in the beginning and end of
             the Features section.
FEATURES             Location/Qualifiers
     source          1. .76922
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="20A7"
                     /clone_lib="Incyte Genomics"
     misc_feature    3861. .3902
                     /note="single clone coverage"
     misc_feature    9095^9096
                     /note="bacterial transposon excised; IS2 element sequence
                     can be found in GenBank Accession Number AE000237.1
                     nucleotides 4452-5782; 1331 bp"
     misc_feature    24082. .24103
                     /note="single clone coverage"
     misc_feature    24907. .25101
                     /note="single clone coverage"
     misc_feature    29277. .29335
                     /note="single clone coverage"
     misc_feature    35088. .35142
                     /note="single clone coverage"
     misc_feature    47307. .47345
                     /note="single clone coverage"
     misc_feature    61397. .61584
                     /note="single clone coverage"
```

misc_feature 62723. .62756
/note="single clone coverage"

BASE COUNT 24246 a 13965 c 13603 g 25108 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 5; Length 76922;
Best Local Similarity 90.0%; Pred. No. 8.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GTCCATTTCCCGTAAATCTT 20
||||||||||||| ||| ||
Db 40811 GTCCATTTCCCGTTAATATT 40792

RESULT 15
AC122615/c
LOCUS AC122615 95594 bp DNA linear HTG 14-JUL-2002
DEFINITION Rattus norvegicus clone CH230-62B1, *** SEQUENCING IN PROGRESS ***, 52 unordered pieces.
ACCESSION AC122615
VERSION AC122615.2 GI:21735366
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 95594)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 95594)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 95594)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jul 12, 2002 this sequence version replaced gi:21205932.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GKBL
Center clone name: CH230-62B1
---------- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 44317 bases at least Q40
Consensus quality: 47172 bases at least Q30
Consensus quality: 48975 bases at least Q20
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 52 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1758: contig of 1758 bp in length
1759	1858: gap of unknown length
1859	2942: contig of 1084 bp in length
2943	3042: gap of unknown length
3043	4588: contig of 1546 bp in length
4589	4688: gap of unknown length
4689	5916: contig of 1228 bp in length
5917	6016: gap of unknown length
6017	7312: contig of 1296 bp in length
7313	7412: gap of unknown length
7413	8429: contig of 1017 bp in length
8430	8529: gap of unknown length
8530	9854: contig of 1325 bp in length
9855	9954: gap of unknown length
9955	10987: contig of 1033 bp in length
10988	11087: gap of unknown length
11088	12415: contig of 1328 bp in length
12416	12515: gap of unknown length
12516	13790: contig of 1275 bp in length
13791	13890: gap of unknown length
13891	15476: contig of 1586 bp in length
15477	15576: gap of unknown length
15577	16654: contig of 1078 bp in length
16655	16754: gap of unknown length
16755	18079: contig of 1325 bp in length
18080	18179: gap of unknown length
18180	19561: contig of 1382 bp in length
19562	19661: gap of unknown length
19662	21000: contig of 1339 bp in length
21001	21100: gap of unknown length
21101	22816: contig of 1716 bp in length
22817	22916: gap of unknown length
22917	24357: contig of 1441 bp in length
24358	24457: gap of unknown length
24458	25571: contig of 1114 bp in length
25572	25671: gap of unknown length
25672	26977: contig of 1306 bp in length

```
  26978    27077: gap of unknown length
  27078    28817: contig of 1740 bp in length
  28818    28917: gap of unknown length
  28918    30270: contig of 1353 bp in length
  30271    30370: gap of unknown length
  30371    31639: contig of 1269 bp in length
  31640    31739: gap of unknown length
  31740    33141: contig of 1402 bp in length
  33142    33241: gap of unknown length
  33242    34943: contig of 1702 bp in length
  34944    35043: gap of unknown length
  35044    36744: contig of 1701 bp in length
  36745    36844: gap of unknown length
  36845    38572: contig of 1728 bp in length
  38573    38672: gap of unknown length
  38673    40895: contig of 2223 bp in length
  40896    40995: gap of unknown length
  40996    43558: contig of 2563 bp in length
  43559    43658: gap of unknown length
  43659    44698: contig of 1040 bp in length
  44699    44798: gap of unknown length
  44799    47265: contig of 2467 bp in length
  47266    47365: gap of unknown length
  47366    48435: contig of 1070 bp in length
  48436    48535: gap of unknown length
  48536    50462: contig of 1927 bp in length
  50463    50562: gap of unknown length
  50563    51711: contig of 1149 bp in length
  51712    51811: gap of unknown length
  51812    52825: contig of 1014 bp in length
  52826    52925: gap of unknown length
  52926    54065: contig of 1140 bp in length
  54066    54165: gap of unknown length
  54166    55170: contig of 1005 bp in length
  55171    55270: gap of unknown length
  55271    56773: contig of 1503 bp in length
  56774    56873: gap of unknown length
  56874    59367: contig of 2494 bp in length
  59368    59467: gap of unknown length
  59468    61502: contig of 2035 bp in length
  61503    61602: gap of unknown length
  61603    62859: contig of 1257 bp in length
  62860    62959: gap of unknown length
  62960    65235: contig of 2276 bp in length
  65236    65335: gap of unknown length
  65336    67974: contig of 2639 bp in length
  67975    68074: gap of unknown length
  68075    70326: contig of 2252 bp in length
  70327    70426: gap of unknown length
  70427    72204: contig of 1778 bp in length
  72205    72304: gap of unknown length
  72305    73946: contig of 1642 bp in length
  73947    74046: gap of unknown length
  74047    77136: contig of 3090 bp in length
  77137    77236: gap of unknown length
  77237    79433: contig of 2197 bp in length
  79434    79533: gap of unknown length
  79534    81647: contig of 2114 bp in length
  81648    81747: gap of unknown length
  81748    83499: contig of 1752 bp in length
  83500    83599: gap of unknown length
  83600    86636: contig of 3037 bp in length
  86637    86736: gap of unknown length
  86737    90821: contig of 4085 bp in length
  90822    90921: gap of unknown length
  90922    95594: contig of 4673 bp in length.

FEATURES             Location/Qualifiers
     source          1. .95594
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"

     Match           84.0%;  Score 16.8;  DB 2;  Length 95594;
     Similarity      90.0%;  Pred. No. 8.6e+02;

     Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy             1 GTCCATTTCCCGTAAATCTT 20
                 |||||||||| ||||||  ||
Db         68221 GTCCATTTCCTGTAAATTTT 68202


Search completed: June 26, 2003, 16:13:28
Job time : 1033.48 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 09:56:18 ; Search time 925.63 Seconds
(without alignments)
565.939 Million cell updates/sec

Title: US-09-355-254F-10
Perfect score: 18
Sequence: 1 agctatgacgttccaagg 18

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	18	100.0	18	6	A89789	A89789 Sequence 11
2	18	100.0	18	6	A90876	A90876 Sequence 11
3	18	100.0	18	6	AX105148	AX105148 Sequence
4	18	100.0	18	6	AX455555	AX455555 Sequence
5	18	100.0	1000	10	S82420S3	S82422 interleukin
6	18	100.0	76047	10	AL669944	AL669944 Mouse DNA
7	18	100.0	110000	2	AC073744_2	Continuation (3 of
8	18	100.0	110000	2	AC073744_3	Continuation (4 of
9	18	100.0	189236	10	AL607030	AL607030 Mouse DNA
c 10	18	100.0	268294	2	AC020885	AC020885 Mus muscu
11	18	100.0	287927	2	AC079530	AC079530 Mus muscu
12	17	94.4	18	6	AX104114	AX104114 Sequence
13	17	94.4	18	6	AX355358	AX355358 Sequence
14	17	94.4	24	6	AX463126	AX463126 Sequence
c 15	17	94.4	24	6	AX463127	AX463127 Sequence
16	16.4	91.1	3184	3	DMPHDOMPT	X97447 D.melanogas
17	16.4	91.1	4663	3	AY122244	AY122244 Drosophil
18	16.4	91.1	70886	2	AC099321	AC099321 Danio rer
19	16.4	91.1	80254	2	AC019531	AC019531 Drosophil
c 20	16.4	91.1	101607	8	AP004334	AP004334 Oryza sat
21	16.4	91.1	102237	2	AC099322	AC099322 Danio rer
c 22	16.4	91.1	152448	8	AP004339	AP004339 Oryza sat
23	16.4	91.1	162726	9	AL157881	AL157881 Human DNA
24	16.4	91.1	169905	9	AL442064	AL442064 Human DNA
c 25	16.4	91.1	186465	2	AC122541	AC122541 Mus muscu
26	16.4	91.1	192681	3	AC011905	AC011905 Drosophil
27	16.4	91.1	252828	2	AL590646	AL590646 Homo sapi
28	16.4	91.1	300829	3	AE003475	AE003475 Drosophil
c 29	16	88.9	157225	2	AC114086	AC114086 Rattus no
c 30	15.4	85.6	1832	3	DVRPPDV	X03936 Drosophila
31	15.4	85.6	2401	3	AB055144	AB055144 Halocynth
32	15.4	85.6	9640	9	AC092955	AC092955 Homo sapi
c 33	15.4	85.6	54623	2	AC098392	AC098392 Rattus no
c 34	15.4	85.6	69507	2	AC079266	AC079266 Mus muscu
35	15.4	85.6	80195	2	AC103059	AC103059 Rattus no
c 36	15.4	85.6	94212	2	AC091848	AC091848 Homo sapi
c 37	15.4	85.6	94296	2	AL160261	AL160261 Homo sapi
38	15.4	85.6	98366	2	AC108548	AC108548 Rattus no
39	15.4	85.6	98775	2	AC121723	AC121723 Rattus no
40	15.4	85.6	107454	2	AC106689	AC106689 Rattus no
41	15.4	85.6	114144	8	U78721	U78721 Arabidopsis
c 42	15.4	85.6	122556	9	AL391863	AL391863 Human DNA
43	15.4	85.6	135203	9	AC025278	AC025278 Homo sapi
c 44	15.4	85.6	140596	2	RN75P15	AL603730 Rattus no
c 45	15.4	85.6	142504	2	AC129671	AC129671 Rattus no

ALIGNMENTS

RESULT 1
A89789
LOCUS A89789 18 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 11 from Patent WO9832462.
ACCESSION A89789
VERSION A89789.1 GI:6738303
KEYWORDS .
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 18)
AUTHORS Lipford,G.B. and Heeg,K.
TITLE PHARMACEUTICAL COMPOSITIONS COMPRISING A POLYNUCLEOTIDE AND OPTIONALLY AN ANTIGEN ESPECIALLY FOR VACCINATION
JOURNAL Patent: WO 9832462-A 11 30-JUL-1998;

```
            LIPFORD GRAYSON B (DE); HEEG KLAUS (DE)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        5 a      4 c      5 g      4 t
ORIGIN

  Query Match           100.0%;  Score 18;  DB 6;  Length 18;
  Best Local Similarity  100.0%;  Pred. No. 7.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATGACGTTCCAAGG 18
              ||||||||||||||||||
Db          1 AGCTATGACGTTCCAAGG 18


RESULT 2
A90876
LOCUS       A90876                   18 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 11 from Patent EP0855184.
ACCESSION   A90876
VERSION     A90876.1  GI:6739275
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Heeg,K.P. and Lipford,G.B.
  TITLE     Pharmaceutical composition comprising a polynucleotide and an
            antigen especially for vaccination
  JOURNAL   Patent: EP 0855184-A 11 29-JUL-1998;
            HEEG KLAUS PROF DR (DE); LIPFORD GRAYSON B DR (DE)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        5 a      4 c      5 g      4 t
ORIGIN

  Query Match           100.0%;  Score 18;  DB 6;  Length 18;
  Best Local Similarity  100.0%;  Pred. No. 7.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATGACGTTCCAAGG 18
              ||||||||||||||||||
Db          1 AGCTATGACGTTCCAAGG 18


RESULT 3
AX105148
LOCUS       AX105148                 18 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 46 from Patent WO0122990.
ACCESSION   AX105148
VERSION     AX105148.1  GI:13921298
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Hartmann,G.D., Bratzler,R.L. and Krieg,A.U.
  TITLE     Methods related to immunostimulatory nucleic acid-induced
            interferon
  JOURNAL   Patent: WO 0122990-A 46 05-APR-2001;
            Coley Pharmaceutical Group, Inc. (US) ; UNIVERSITY OF IOWA RESEARCH
            FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic Oligonucleotide"
BASE COUNT        5 a      4 c      5 g      4 t
ORIGIN

  Query Match           100.0%;  Score 18;  DB 6;  Length 18;
  Best Local Similarity  100.0%;  Pred. No. 7.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATGACGTTCCAAGG 18
              ||||||||||||||||||
Db          1 AGCTATGACGTTCCAAGG 18


RESULT 4
AX455555
LOCUS       AX455555                 18 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 32 from Patent WO0222809.
ACCESSION   AX455555
VERSION     AX455555.1  GI:21714623
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bauer,S., Lipford,G. and Wagner,H.
  TITLE     Process for high throughput screening of cpg-based
            immuno-agonist/antagonist
  JOURNAL   Patent: WO 0222809-A 32 21-MAR-2002;
            Coley Pharmaceutical GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic oligonucleotide"
BASE COUNT        5 a      4 c      5 g      4 t
ORIGIN

  Query Match           100.0%;  Score 18;  DB 6;  Length 18;
  Best Local Similarity  100.0%;  Pred. No. 7.4;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATGACGTTCCAAGG 18
              ||||||||||||||||||
Db          1 AGCTATGACGTTCCAAGG 18


RESULT 5
S82420S3
LOCUS       S82420S3                 1000 bp    DNA     linear   ROD 03-DEC-1996
DEFINITION  interleukin-12 p40 subunit [mice, Genomic, 1000 nt, segment 3 of
            7].
ACCESSION   S82422
VERSION     S82422.1  GI:1699185
KEYWORDS    .
SEGMENT     3 of 7
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1000)
  AUTHORS   Tone,Y., Thompson,S.A., Babik,J.M., Nolan,K.F., Tone,M., Raven,C.
            and Waldmann,H.
  TITLE     Structure and chromosomal location of the mouse interleukin-12 p35
            and p40 subunit genes
  JOURNAL   Eur. J. Immunol. 26 (6), 1222-1227 (1996)
  MEDLINE   96257799
   PUBMED   8647196
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 178372] from the original journal article.
            This sequence comes from Fig. 4.
            Map location: 11.
FEATURES             Location/Qualifiers
     source          1. .1000
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
BASE COUNT      314 a    221 c    207 g    258 t
```

ICC

[redacted]:14:36 2003

IN

Query Match 100.0%; Score 18; DB 10; Length 1000;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCTATGACGTTCCAAGG 18
||||||||||||||||||
Db 28 AGCTATGACGTTCCAAGG 45

RESULT 6
AL669944
LOCUS AL669944 76047 bp DNA linear ROD 05-APR-2002
DEFINITION Mouse DNA sequence from clone RP23-388G23 on chromosome 11, complete sequence.
ACCESSION AL669944
VERSION AL669944.8 GI:20068715
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Dunn,M.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Apr 8, 2002 this sequence version replaced gi:19309600.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-388G23 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES Location/Qualifiers
source 1. .76047
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-388G23"
/clone_lib="RPCI-23"

BASE COUNT 22455 a 16245 c 16275 g 21072 t
ORIGIN

Query Match 100.0%; Score 18; DB 10; Length 76047;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCTATGACGTTCCAAGG 18
||||||||||||||||||
Db 913 AGCTATGACGTTCCAAGG 930

RESULT 7
AC073744_2
WPCOMMENT LOCUS AC073744 Accession AC073744
Sequence split into 5 fragments
Fragment Name Begin End
AC073744_0 1 110000
AC073744_1 100001 210000
AC073744_2 200001 310000
AC073744_3 300001 410000
AC073744_4 400001 466859
Continuation (3 of 5) of AC073744 from base 200001 (AC073744 Mus musculus clone RP23-

Query Match 100.0%; Score 18; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCTATGACGTTCCAAGG 18
||||||||||||||||||
Db 105590 AGCTATGACGTTCCAAGG 105607

RESULT 8
AC073744_3
WPCOMMENT LOCUS AC073744 Accession AC073744
Sequence split into 5 fragments
Fragment Name Begin End
AC073744_0 1 110000
AC073744_1 100001 210000
AC073744_2 200001 310000
AC073744_3 300001 410000
AC073744_4 400001 466859
Continuation (4 of 5) of AC073744 from base 300001 (AC073744 Mus musculus clone RP2

Query Match 100.0%; Score 18; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCTATGACGTTCCAAGG 18
||||||||||||||||||
Db 5590 AGCTATGACGTTCCAAGG 5607

RESULT 9
AL607030
LOCUS AL607030 189236 bp DNA linear ROD 05-APR-2002
DEFINITION Mouse DNA sequence from clone RP23-304O12 on chromosome 11, complete sequence.
ACCESSION AL607030
VERSION AL607030.15 GI:19847865
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Brown,A.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Mar 29, 2002 this sequence version replaced gi:19571965.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers giv

in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; WP:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-304O12 is from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

```
FEATURES             Location/Qualifiers
     source          1. .189236
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="11"
                     /clone="RP23-304O12"
                     /clone_lib="RPCI-23"
BASE COUNT     55297 a  41962 c  40531 g  51446 t
ORIGIN
```

```
  Query Match                 100.0%;  Score 18;  DB 10;  Length 189236;
  Best Local Similarity   100.0%;  Pred. No. 9.8;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATGACGTTCCAAGG 18
              ||||||||||||||||||
Db     188148 AGCTATGACGTTCCAAGG 188165
```

```
RESULT 10
AC020885/c
LOCUS       AC020885               268294 bp    DNA     linear   HTG 16-FEB-2000
DEFINITION  Mus musculus clone RP23-464I1, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC020885
VERSION     AC020885.2  GI:6980212
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 268294)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Mouse
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 268294)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 16, 2000 this sequence version replaced gi:6686423.
            * NOTE: This record contains 183 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1     1013: contig of 1013 bp in length
            *     1014     1615: gap of unknown length
            *                    contig of 602 bp in length
            *     1616     2405: gap of unknown length
            *                    contig of 790 bp in length
            *     2406     3234: gap of unknown length
            *                    contig of 829 bp in length
            *     3235     3533: gap of unknown length
            *                    contig of 299 bp in length
            *     3534     4467: gap of unknown length
            *                    contig of 934 bp in length
            *     4468     5401: gap of unknown length
            *                    contig of 934 bp in length
            *     5402     6154: gap of unknown length
            *                    contig of 753 bp in length
            *     6155     6629: gap of unknown length
            *                    contig of 475 bp in length
            *     6630     7176: gap of unknown length
            *                    contig of 547 bp in length
            *     7177     7824: gap of unknown length
            *                    contig of 648 bp in length
            *     7825     8656: gap of unknown length
            *                    contig of 832 bp in length
            *     8657     8865: gap of unknown length
            *                    contig of 209 bp in length
            *     8866     9544: gap of unknown length
            *                    contig of 679 bp in length
            *     9545    10280: gap of unknown length
            *                    contig of 736 bp in length
            *    10281    10975: gap of unknown length
            *                    contig of 695 bp in length
            *    10976    11742: gap of unknown length
            *                    contig of 767 bp in length
            *    11743    12178: gap of unknown length
            *                    contig of 436 bp in length
            *    12179    12887: gap of unknown length
            *                    contig of 709 bp in length
            *    12888    13607: gap of unknown length
            *                    contig of 720 bp in length
            *    13608    13911: gap of unknown length
            *                    contig of 304 bp in length
            *    13912    14819: gap of unknown length
            *                    contig of 908 bp in length
            *    14820    15571: gap of unknown length
            *                    contig of 752 bp in length
            *    15572    15872: gap of unknown length
            *                    contig of 301 bp in length
            *    15873    16608: gap of unknown length
            *                    contig of 736 bp in length
            *    16609    16851: gap of unknown length
            *                    contig of 243 bp in length
            *    16852    17277: gap of unknown length
            *                    contig of 426 bp in length
            *    17278    18057: gap of unknown length
            *                    contig of 780 bp in length
            *    18058    18704: gap of unknown length
            *                    contig of 647 bp in length
            *    18705    19414: gap of unknown length
            *                    contig of 710 bp in length
            *    19415    20372: gap of unknown length
            *                    contig of 958 bp in length
            *    20373    21010: gap of unknown length
            *                    contig of 638 bp in length
            *    21011    21258: gap of unknown length
            *                    contig of 248 bp in length
            *    21259    22363: gap of unknown length
            *                    contig of 1105 bp in length
            *    22364    23639: gap of unknown length
            *                    contig of 1276 bp in length
            *    23640    24853: gap of unknown length
            *                    contig of 1214 bp in length
            *    24854    25795: gap of unknown length
            *                    contig of 942 bp in length
            *    25796    27027: gap of unknown length
            *                    contig of 1232 bp in length
            *    27028    27729: gap of unknown length
            *                    contig of 702 bp in length
            *    27730    28722: gap of unknown length
            *                    contig of 993 bp in length
            *    28723    29714: gap of unknown length
            *                    contig of 992 bp in length
            *    29715    30858: gap of unknown length
            *                    contig of 1144 bp in length
            *    30859    31397: gap of unknown length
            *                    contig of 539 bp in length
            *    31398    31671: gap of unknown length
                                 contig of 274 bp in length
```

```
                gap of unknown length
     31672  32959: contig of 1288 bp in length
                gap of unknown length
     32960  33791: contig of 832 bp in length
                gap of unknown length
     33792  34848: contig of 1057 bp in length
                gap of unknown length
     34849  35060: contig of 212 bp in length
                gap of unknown length
     35061  35776: contig of 716 bp in length
                gap of unknown length
     35777  35903: contig of 127 bp in length
                gap of unknown length
     35904  36602: contig of 699 bp in length
                gap of unknown length
     36603  37110: contig of 508 bp in length
                gap of unknown length
     37111  38059: contig of 949 bp in length
                gap of unknown length
     38060  38730: contig of 671 bp in length
                gap of unknown length
     38731  39790: contig of 1060 bp in length
                gap of unknown length
     39791  40327: contig of 537 bp in length
                gap of unknown length
     40328  40442: contig of 115 bp in length
                gap of unknown length
     40443  41688: contig of 1246 bp in length
                gap of unknown length
     41689  42688: contig of 1000 bp in length
                gap of unknown length
     42689  43847: contig of 1159 bp in length
                gap of unknown length
     43848  44511: contig of 664 bp in length
                gap of unknown length
     44512  45780: contig of 1269 bp in length
                gap of unknown length
     45781  47202: contig of 1422 bp in length
                gap of unknown length
     47203  48647: contig of 1445 bp in length
                gap of unknown length
     48648  49652: contig of 1005 bp in length
                gap of unknown length
     49653  50485: contig of 833 bp in length
                gap of unknown length
     50486  51527: contig of 1042 bp in length
                gap of unknown length
     51528  53119: contig of 1592 bp in length
                gap of unknown length
     53120  53623: contig of 504 bp in length
                gap of unknown length
     53624  54669: contig of 1046 bp in length
                gap of unknown length
     54670  55309: contig of 640 bp in length
                gap of unknown length
     55310  56382: contig of 1073 bp in length
                gap of unknown length
     56383  56625: contig of 243 bp in length
                gap of unknown length
     56626  57345: contig of 720 bp in length
                gap of unknown length
     57346  57505: contig of 160 bp in length
                gap of unknown length
     57506  58908: contig of 1403 bp in length
                gap of unknown length
     58909  59968: contig of 1060 bp in length
                gap of unknown length
     59969  61377: contig of 1409 bp in length
                gap of unknown length
     61378  62291: contig of 914 bp in length
                gap of unknown length
     62292  63319: contig of 1028 bp in length
                gap of unknown length
     63320  63818: contig of 499 bp in length
                gap of unknown length
     63819  64823: contig of 1005 bp in length
                gap of unknown length
     64824  65269: contig of 446 bp in length
                gap of unknown length
     65270  65904: contig of 635 bp in length
                gap of unknown length
     65905  66815: contig of 911 bp in length
                gap of unknown length

Query Match          100.0%;  Score 18;  DB 2;  Length 268294;
Best Local Similarity  100.0%;  Pred. No. 9.9;
Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATGACGTTCCAAGG 18
              ||||||||||||||||||
Db     217636 AGCTATGACGTTCCAAGG 217619


RESULT 11
AC079530
LOCUS       AC079530               287927 bp    DNA     linear   HTG 02-SEP-2000
DEFINITION  Mus musculus clone RP23-342G15, WORKING DRAFT SEQUENCE, 53
            unordered pieces.
ACCESSION   AC079530
VERSION     AC079530.1  GI:9964895
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 287927)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Mouse
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 287927)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 1868359
            Center clone name: RPCI-23_342G15
            -------
            Summary Statistics
            Consensus quality: 236925 bases at least Q40
            Consensus quality: 254207 bases at least Q30
            Consensus quality: 258663 bases at least Q20
            Estimated insert size: 210000; pulse field gel estimation
            Estimated insert size: 282727; sum-of-contigs estimation
            Quality coverage:  14.02 in Q20 bases; pulse field gel estimation
            Quality coverage:  10.41 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 53 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1636: contig of 1636 bp in length
            *     1637     1736: gap of unknown length
            *     1737     3079: contig of 1343 bp in length
            *     3080     3179: gap of unknown length
            *     3180     4715: contig of 1536 bp in length
            *     4716     4815: gap of unknown length
            *     4816     5816: contig of 1001 bp in length
```

```
*     5817    5916: gap of unknown length
*     5917    7443: contig of 1527 bp in length
*     7444    7543: gap of unknown length
*     7544    9268: contig of 1725 bp in length
*     9269    9368: gap of unknown length
*     9369   10739: contig of 1371 bp in length
*    10740   10839: gap of unknown length
*    10840   12239: contig of 1400 bp in length
*    12240   12339: gap of unknown length
*    12340   14093: contig of 1754 bp in length
*    14094   14193: gap of unknown length
*    14194   15523: contig of 1330 bp in length
*    15524   15623: gap of unknown length
*    15624   17085: contig of 1462 bp in length
*    17086   17185: gap of unknown length
*    17186   18230: contig of 1045 bp in length
*    18231   18330: gap of unknown length
*    18331   19523: contig of 1193 bp in length
*    19524   19623: gap of unknown length
*    19624   20712: contig of 1089 bp in length
*    20713   20812: gap of unknown length
*    20813   21950: contig of 1138 bp in length
*    21951   22050: gap of unknown length
*    22051   23824: contig of 1774 bp in length
*    23825   23924: gap of unknown length
*    23925   25225: contig of 1301 bp in length
*    25226   25325: gap of unknown length
*    25326   26530: contig of 1205 bp in length
*    26531   26630: gap of unknown length
*    26631   28598: contig of 1968 bp in length
*    28599   28698: gap of unknown length
*    28699   29722: contig of 1024 bp in length
*    29723   29822: gap of unknown length
*    29823   30922: contig of 1100 bp in length
*    30923   31022: gap of unknown length
*    31023   32970: contig of 1948 bp in length
*    32971   33070: gap of unknown length
*    33071   35474: contig of 2404 bp in length
*    35475   35574: gap of unknown length
*    35575   36870: contig of 1296 bp in length
*    36871   36970: gap of unknown length
*    36971   38093: contig of 1123 bp in length
*    38094   38193: gap of unknown length
*    38194   40123: contig of 1930 bp in length
*    40124   40223: gap of unknown length
*    40224   41625: contig of 1402 bp in length
*    41626   41725: gap of unknown length
*    41726   46312: contig of 4587 bp in length
*    46313   46412: gap of unknown length
*    46413   50230: contig of 3818 bp in length
*    50231   50330: gap of unknown length
*    50331   51665: contig of 1335 bp in length
*    51666   51765: gap of unknown length
*    51766   53285: contig of 1520 bp in length
*    53286   53385: gap of unknown length
*    53386   56499: contig of 3114 bp in length
*    56500   56599: gap of unknown length
*    56600   57786: contig of 1187 bp in length
*    57787   57886: gap of unknown length
*    57887   63353: contig of 5467 bp in length
*    63354   63453: gap of unknown length
*    63454   69163: contig of 5710 bp in length
*    69164   69263: gap of unknown length
*    69264   74152: contig of 4889 bp in length
*    74153   74252: gap of unknown length
*    74253   79702: contig of 5450 bp in length
*    79703   79802: gap of unknown length
*    79803   86526: contig of 6724 bp in length
*    86527   86626: gap of unknown length
*    86627   92661: contig of 6035 bp in length
*    92662   92761: gap of unknown length
*    92762   97809: contig of 5048 bp in length
*    97810   97909: gap of unknown length
*    97910   102775: contig of 4866 bp in length
*   102776   102875: gap of unknown length
*   102876   108736: contig of 5861 bp in length
*   108737   108836: gap of unknown length
*   108837   117122: contig of 8286 bp in length
*   117123   117222: gap of unknown length
*   117223   124266: contig of 7044 bp in length
*   124267   124366: gap of unknown length
*   124367   137828: contig of 13462 bp in length
*   137829   137928: gap of unknown length
*   137929   144300: contig of 6372 bp in length
*   144301   144400: gap of unknown length
*   144401   159887: contig of 15487 bp in length
*   159888   159987: gap of unknown length
*   159988   170526: contig of 10539 bp in length
*   170527   170626: gap of unknown length
*   170627   185958: contig of 15332 bp in length
*   185959   186058: gap of unknown length
*   186059   203429: contig of 17371 bp in length
*   203430   203529: gap of unknown length
*   203530   222635: contig of 19106 bp in length
*   222636   222735: gap of unknown length
*   222736   242871: contig of 20136 bp in length
*   242872   242971: gap of unknown length
*   242972   287927: contig of 44956 bp in length.
FEATURES             Location/Qualifiers
     source          1. .287927
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="RP23-342G15"
                     /clone_lib="RPCI mouse BAC library 23"
BASE COUNT     81254 a  63612 c  61615 g  76140 t   5306 others
ORIGIN

  Query Match                 100.0%;  Score 18;  DB 2;  Length 287927;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCTATGACGTTCCAAGG 18
              ||||||||||||||||||
Db     174744 AGCTATGACGTTCCAAGG 174761


RESULT 12
AX104114
LOCUS       AX104114                  18 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 306 from Patent WO0122972.
ACCESSION   AX104114
VERSION     AX104114.1  GI:13920311
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 306 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT         4 a       4 c       6 g       4 t
ORIGIN

  Query Match                 94.4%;  Score 17;  DB 6;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GCTATGACGTTCCAAGG 18
              |||||||||||||||||
Db          1 GCTATGACGTTCCAAGG 17
```

```
RESULT 13
AX355358
LOCUS       AX355358                  18 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 386 from Patent WO0197843.
ACCESSION   AX355358
VERSION     AX355358.1  GI:18620026
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Weiner,G. and Hartmann,G.
  TITLE     Methods for enhancing antibody-induced cell lysis and treating
            cancer
  JOURNAL   Patent: WO 0197843-A 386 27-DEC-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic oligonucleotide-phosphorothioate
                     backbone"
BASE COUNT        4 a      4 c      6 g      4 t
ORIGIN

  Query Match           94.4%;  Score 17;  DB 6;  Length 18;
  Best Local Similarity  100.0%;  Pred. No. 31;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GCTATGACGTTCCAAGG 18
              |||||||||||||||||
Db          1 GCTATGACGTTCCAAGG 17


RESULT 14
AX463126
LOCUS       AX463126                  24 bp    DNA     linear   PAT 15-JUL-2002
DEFINITION  Sequence 9 from Patent WO0250108.
ACCESSION   AX463126
VERSION     AX463126.1  GI:21886107
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Marchal,G., Pescher,P. and Romain,F.
  TITLE     Immunogenic glycopeptides, screening, preparation and uses
  JOURNAL   Patent: WO 0250108-A 9 27-JUN-2002;
            PASTEUR INSTITUT (FR)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="linker"
BASE COUNT        5 a      8 c      7 g      4 t
ORIGIN

  Query Match           94.4%;  Score 17;  DB 6;  Length 24;
  Best Local Similarity  100.0%;  Pred. No. 31;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GCTATGACGTTCCAAGG 18
              |||||||||||||||||
Db          4 GCTATGACGTTCCAAGG 20


RESULT 15
AX463127/c
LOCUS       AX463127                  24 bp    DNA     linear   PAT 15-JUL-2002
DEFINITION  Sequence 10 from Patent WO0250108.
ACCESSION   AX463127
VERSION     AX463127.1  GI:21886108
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Marchal,G., Pescher,P. and Romain,F.
  TITLE     Immunogenic glycopeptides, screening, preparation and uses
  JOURNAL   Patent: WO 0250108-A 10 27-JUN-2002;
            PASTEUR INSTITUT (FR)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="linker"
BASE COUNT        4 a      7 c      8 g      5 t
ORIGIN

  Query Match           94.4%;  Score 17;  DB 6;  Length 24;
  Best Local Similarity  100.0%;  Pred. No. 31;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GCTATGACGTTCCAAGG 18
              |||||||||||||||||
Db         21 GCTATGACGTTCCAAGG 5


Search completed: June 26, 2003, 16:13:10
Job time : 929.63 secs
```